VOLUME 37 MARCH 1959 NUMBER 3

Canadian Journal of Physics

Editor: H. E. DUCKWORTH

Associate Editors:

L. G. ELLIOTT, *Atomic Energy of Canada, Ltd., Chalk River*
J. S. FOSTER, *McGill University*
G. HERZBERG, *National Research Council of Canada*
L. LEPRINCE-RINGUET, *Ecole Polytechnique, Paris*
B. W. SARGENT, *Queen's University*
G. M. VOLKOFF, *University of British Columbia*
W. H. WATSON, *University of Toronto*
G. A. WOONTON, *McGill University*

Published by THE NATIONAL RESEARCH COUNCIL

OTTAWA CANADA

CANADIAN JOURNAL OF PHYSICS

Under the authority of the Chairman of the Committee of the Privy Council on Scientific and Industrial Research, the National Research Council issues THE CANADIAN JOURNAL OF PHYSICS and five other journals devoted to the publication, in English or French, of the results of original scientific research. Matters of general policy concerning these journals are the responsibility of a joint Editorial Board consisting of: members representing the National Research Council of Canada; the Editors of the Journals; and members representing the Royal Society of Canada and four other scientific societies.

EDITORIAL BOARD

Representatives of the National Research Council

A. Gauthier, *University of Montreal*
R. B. Miller, *University of Alberta*
H. G. Thode, *McMaster University*
D. L. Thomson, *McGill University*

Editors of the Journals

D. L. Bailey, *University of Toronto*
T. W. M. Cameron, *Macdonald College*
H. E. Duckworth, *McMaster University*
K. A. C. Elliott, *Montreal Neurological Institute*
Léo Marion, *National Research Council*
R. G. E. Murray, *University of Western Ontario*

Representatives of Societies

D. L. Bailey, *University of Toronto*
Royal Society of Canada
T. W. M. Cameron, *Macdonald College*
Royal Society of Canada
H. E. Duckworth, *McMaster University*
Royal Society of Canada
Canadian Association of Physicists
K. A. C. Elliott, *Montreal Neurological Institute*
Canadian Physiological Society
P. R. Gendron, *University of Ottawa*
Chemical Institute of Canada
R. G. E. Murray, *University of Western Ontario*
Canadian Society of Microbiologists
T. Thorvaldson, *University of Saskatchewan*
Royal Society of Canada

Ex officio

Léo Marion (Editor-in-Chief), *National Research Council*
J. B. Marshall (Administration and Awards), *National Research Council*

Manuscripts for publication should be submitted to Dr. H. E. Duckworth, Editor, Canadian Journal of Physics, Hamilton College, McMaster University, Hamilton, Ontario.
For instructions on preparation of copy, see **NOTES TO CONTRIBUTORS** (back cover).

Proof, correspondence concerning proof, and orders for reprints should be sent to the Manager, Editorial Office (Research Journals), Division of Administration and Awards, National Research Council, Ottawa 2, Canada.

Subscriptions, renewals, requests for single or back numbers, and all remittances should be sent to Division of Administration and Awards, National Research Council, Ottawa 2, Canada. Remittances should be made payable to the Receiver General of Canada, credit National Research Council.

The journals published, frequency of publication, and prices are:

Canadian Journal of Biochemistry and Physiology	Monthly	$9.00 a year
Canadian Journal of Botany	Bimonthly	$6.00 a year
Canadian Journal of Chemistry	Monthly	$12.00 a year
Canadian Journal of Microbiology	Bimonthly	$6.00 a year
Canadian Journal of Physics	Monthly	$9.00 a year
Canadian Journal of Zoology	Bimonthly	$5.00 a year

The price of regular single numbers of all journals is $2.00.

Canadian Journal of Physics

Issued by THE NATIONAL RESEARCH COUNCIL OF CANADA

VOLUME 37 MARCH 1959 NUMBER 3

A PROPOSED MECHANISM FOR ICE FRICTION[1]

C. D. NIVEN

ABSTRACT

The collapsible structure of ice is stressed. The adhesion bond is visualized as being associated with a readjustment of a partially collapsed or distorted molecular structure when cooling sets in after plastic flow has occurred in the asperity. With this general idea in view the friction mechanism on ice is compared in detail with that on a non-collapsible material, and explanations for observations are suggested.

INTRODUCTION

In previous communications (Niven 1954, 1956, 1956, 1958) the results of some measurements of ice friction carried out in the cold room have been described. These investigations were undertaken in order to acquire some information on obscure points, such as the effect of high loading, the effect of heating the runner, the effect of speed, the behavior of rubber. In the present communication an attempt is made to fit the results into the modern adhesion theory of friction which explains the behavior of metals.

Up to the present time two ideas have dominated the thinking of engineers and scientists in connection with ice friction: one idea is that pressure melting supplies a film of water on the surface: the other idea is that frictional heating supplies that film (Bowden and Hughes 1939). The two ideas have a point in common, namely, that if melting can be accounted for, the low friction of ice is explained and the problem is solved. Researchers have accepted this assumption uncritically: for instance Klein (1950) tried to see the water by using a window on the ski: Bowden and Hughes (1939) used an electrical method to prove the presence of water on the ice surface: McConica (1951) suggested the possibility of the adsorption of water molecules or "hydrol" by the ice: Weyl (1951) suggested that the surface of the ice was liquid owing to polarity adjustments in the layers beneath.

But to merely account for a film of water on the surface of the ice is far too superficial an approach to the ice friction problem because apart altogether from the low friction value, ice friction violates Amontons' law at high

[1]Manuscript received November 24, 1958.
Contribution from the Division of Applied Physics, National Research Council, Ottawa, Canada.
Issued as N.R.C. No. 5064.

loading: it violates very seriously the approximate law that friction is independent of temperature; and it violates to a small extent the other approximate friction law that friction is independent of speed—particularly at low speeds. Before a substance stands out as being so exceptional there must be something fundamentally different about it. Any attempt therefore to explain ice friction entirely by a frictional heating mechanism is unsatisfactory because this implies that all hard substances which melt should have ice-like friction qualities just below their melting points. But the following more specific objection to the frictional heating explanation was noted by McConica (1950): a corollary of that explanation is that sliders with high thermal conductivity should have high friction: McConica points out that magnesium, which is a good heat conductor, has a low friction.

The other and older theory—the pressure melting idea of Reynolds—has been criticized and practically rejected because the pressures necessary to cause liquefaction are very high. However, this theory has a point to recommend it, in that it picks out a unique quality of water—namely its expansion on freezing—and makes the theory depend on this peculiarity: indeed, some knowledgeable scientists still cling to this theory on the grounds that the pressure on the tip of an asperity is great enough to liquefy a tiny quantity of water: but this implies that the area of real contact is very small indeed so that this explanation is not satisfactory.

Wide differences in the friction values of the various substances on ice suggest that one must go beyond the macroscopic physical ideas of melting by pressure or heat and imagine what may be going on at the surface of the ice on an atomic or molecular scale—in other words, apply to this problem some information available about the structure of water and ice.

THE EFFECT OF HIGH LOADING AND OF HEATED RUNNERS ON ICE FRICTION

The writer's attention was drawn to the effect of high loading by Mr. W. E. Wakefield of the Forest Products Laboratory in Ottawa, who stated that after a sleigh was very heavily loaded the drag increased very little if the load was increased. In his report Wakefield (1938) merely emphasized that friction drops at high loading, which is really an understatement of a very remarkable phenomenon, because in mild winter temperatures laboratory tests (Niven 1954) showed that drag was practically independent of load at high loadings: in severe winter temperatures the results indicated that far higher loads would be required for the drag to be constant with load. Figure 1 illustrates these findings and simply forces one to the conclusion that the pressure on the ice cannot be ignored in any sound explanation of ice friction. Amontons' law—virtually the king-pin of friction theory for ordinary substances—does not hold good for ice if the pressure is high and the temperature moderate.

The work on the heated runners was carried out to see if they had a lower friction on ice than unheated runners. If a film of water on the ice surface was the explanation of low friction, it was expected that heating the runner

would assist the formation of the film and so reduce the friction. Pfalzner (1947) and the writer (Niven 1956) were independently forced to the same conclusion—namely that heating the runner reduced the friction by only a very slight amount. The experiments were disappointing but the fact that a slight reduction could be observed under certain circumstances e.g. at lower temperatures, rather supported than refuted the frictional heating explanation of ice friction.

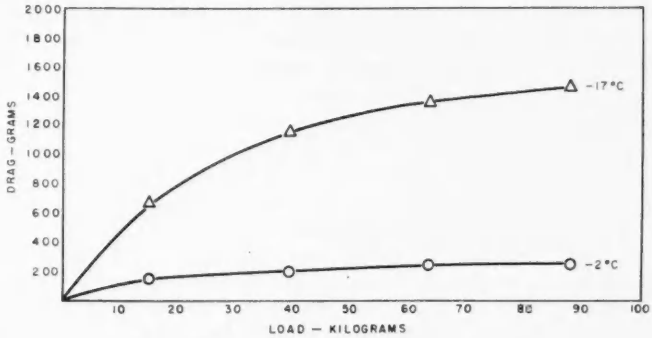


FIG. 1. Drag vs. load curves at −2° C and −17° C for a miniature stainless steel sleigh with an apparent contact area of 1.6 sq. cm.

These tests at high loading and with heated runners suggested that if the explanation of ice friction were merely a problem of accounting for a film of water, the pressure-melting idea and the frictional-heat idea were both partly right. Unfortunately the basic idea of lubrication by means of the film of water seems open to criticism, because water is certainly not a good lubricant in general and in contact with a cold surface it is an adhesive. The idea that the liquid film reduces friction seems to have crept in because the idea of molecules gliding past each other in the liquid state was quite familiar. If for some reason ice happened to be a solid in which the molecules could turn slightly or swivel in groups then friction on such a solid might be low regardless of the presence of the liquid film. However, in the meanwhile let us consider the effect of a liquid film in contact with the cold surface.

The film acts as an adhesive and therefore anything which tends to produce it also tends to create a condition whereby the slider can adhere to the surface. The water film, if it does exist at all, must then perform two functions in the friction mechanism—firstly, it must reduce friction by creating a condition whereby molecules can easily slip past each other and secondly, it must augment friction by creating a condition whereby one surface can adhere to the other. It is these two conflicting tendencies which complicate the explanation of ice friction: but before pursuing this point farther, the sliding process of metals which obey the normal friction laws must be very briefly considered.

THE SLIDING FRICTION MECHANISM FOR NORMAL METALS COMPARED WITH THE MECHANISM FOR ICE

Sliding friction is mainly an adhesion phenomenon with a certain amount of ploughing action in the surface. Pressure and rubbing cause the tips of the asperities to weld together in the case of the normal metals and the force required to shear these welds is what accounts for the frictional force according to modern theory. When the two surfaces are not both metallic it seems necessary to assume other kinds of bonds than the metallic bond: the strength and type of these bonds presumably depend on the electronic structures of the surface atoms.

To fully appreciate what is happening during the sliding mechanism it is well to break up this general description of sliding friction into five events, which must be considered to blend into each other as the one asperity passes the other.

The first event is a collision between an asperity projecting down from the slider and an asperity projecting up from the surface, resulting in a drastic pushing of molecules in one or both asperities, accompanied, of course, by a stretching and straining of bonds—in fact, a plastic flow. This amounts to saying that heat is produced.

The second event in the mechanism is that the two colliding asperities take a greater share of the load on the slider and, over the little area of real contact formed by the plastic flow, molecules of the surface are brought into close contact with molecules of the slider—pressed against them and rubbed by them.

The third event is that the surfaces adhere—in the case of metals, weld together.

The fourth event is that plastic flow occurs in one or both of the asperities as the forward movement of the slider causes the one asperity to drag the other along with it.

The fifth event is that the asperities separate possibly with a piece of one of them torn out and sticking on to the other.

If the main part of the frictional force is accounted for by the shearing of the welds, this means that frictional heating is to be accounted for by the plastic flow and breaking of bonds after the welding process. However, since Amontons' law for ordinary metals is explained by the dependence of the area of real contact on the load for materials which deform plastically and since the weight is continually being carried by new asperities as the slider moves forward, some plastic flow and therefore heating occurs before welding: therefore there is expenditure of energy before as well as after welding, or in other words part of the frictional force is employed in the pushing of one asperity by the other. This apparently small detail is of considerable importance in explaining ice friction.

The structure of ice is an open structure (Bernal and Fowler 1933) consisting of an atom of oxygen surrounded by four other oxygen atoms held to the central atom by hydrogen bonds and forming a tetrahedron around it. This structure differs basically from another well-known tetrahedral structure

often applying to metals, namely the close-packed structure of the hexagonal or face-centered cubic types, where the central atom is missing and the four atoms forming the tetrahedron nest snugly into each other.

To appreciate then the difference between ice friction and the friction of quite a number of ordinary metals, one must compare an asperity composed of atoms or molecules in the open tetrahedral structure with an asperity having the close-packed structure. When the *close-packed* structure is hit by a slider asperity, atomic disarrangement occurs and this may be regarded as a tendency to *melt* the surface asperity: pressure, however, tends to keep the molecules nested in place and this may be regarded as a tendency to *solidify* it. On the other hand when the *ice* asperity is hit by a slider asperity the molecular disarrangement which occurs is accompanied by a tendency to make the open structure collapse and get into a structure more like that of the water structure: this amounts to saying that when the open-structured ice asperity is hit, agitation and pressure co-operate in tending to melt the asperity. With the close-packed asperity they work against each other and welding, instead of the tendency to melt, results.

If the load on the slider were so great that agitation and pressure completely liquefied the asperity then no further increase in pressure should cause any more melting, and drag should remain constant. This condition would most easily arise at temperatures near to 0° C when the structure should collapse more easily than in severe temperatures (see Fig. 1).

Since water retains its tetrahedral structure after melting occurs, a more accurate description of the pressure-melting mechanism is a swivelling of open tetrahedra held together by electrostatic bonds. When the pressure is large enough to completely "crush" the ice structure into the liquid tetrahedral structure, no further pressure—within reasonable limits—can do any more crushing: for some reason this structure is mobile. This, of course, was essentially the point which Reynolds intuitively perceived long before the tetrahedral structure of water was heard of. The present explanation of ice friction assumes that in between the completely "crushed" structure applicable to liquid water and the uncrushed or unloaded structure applicable to ice there is a continuous series of structures in which the tetrahedra have swivelled to a greater or less extent depending on the load. That is why throughout the above discussion the formation of water of lubrication has been referred to conditionally. The process during the friction mechanism which is usually referred to as the formation of water of lubrication is visualized by the writer as a swivelling of tetrahedra to form a structure more like the water structure—a *tendency* to liquefaction rather than a complete liquefaction. The suggestion is that this partial swivelling of tetrahedra—somewhat reminiscent of a ball moving in a ball race—reduces the friction without the actual production of a water film as a result of complete swivelling.

Returning now to the general friction mechanism for normal metals, after the agitation due to collision of the two asperities has taken place the slider asperity starts to leave the surface asperity. At this stage the welds which have been formed by agitation and pressure start to shear. But with ice the

situation is quite different because the welds have not been formed: in fact, the slider asperity must be visualized as being in contact with a spot of ice surface which is to be regarded as having a distorted tetrahedral structure more applicable to water than the undistorted ice structure was. The surrounding temperature is below 0° C and so "freezing" begins, that is to say the distorted structure returns to the normal ice structure. It is at this juncture that the question of whether or not the slider is composed of a material to which ice will adhere becomes highly important, but before discussing the adhesion bond, let us continue to concentrate attention on the freezing process which hardens up the ice asperity immediately after it has been subjected to molecular disarrangement and pressure.

If the temperature of the main body of the ice and of the slider is well below 0° C the freezing process will be rapid and the bond will be strong. Therefore ice friction should be highly dependent on temperature—this agrees with observation. The rapidity with which the asperity will harden up should be proportional to the temperature difference of the asperity and the ice: in other words the contribution to the friction made by the adhesion bond should depend on the number of degrees the ice temperature is below 0° C.

Furthermore since the formation of the adhesion bond is due to heat loss by conduction, it is to be expected that the bond will be poor if the time during which heat loss can take place is short, that is to say if the slider is moving quickly. Observation (Niven 1956) shows that ice friction drops slightly with increase of speed—particularly at low speeds.

The mechanism for ice friction described above provides then for the observations that ice disobeys all the laws of friction—Amontons' law at high loading and also the two approximate laws regarding independence on temperature and speed.

For the sake of clarity only a close-packed structure was discussed above in comparing the normal friction mechanism with the ice friction mechanism but the argument applies not only to a close-packed structure but to any structure whose molecules tend to nest closely into each other and keep their arrangement when compressed. This, incidentally, amounts to saying that the argument does not apply to those substances which expand on freezing, and so the suggestion is that any solid which expands on freezing should have friction qualities to some extent like those of ice and different from normal substances. A hint that this conception may be correct is an observation by Bowden and Freitag (1955) that bismuth has a low friction coefficient comparable with that of ice.

THE ADHESION BOND

When metal slides on metal the adhesion bond must be a metallic bond because Bowden (1950) and his collaborators have shown that the bonds of the weld can be as strong as the bonds which hold the atoms together in the metal. The adhesion bond when ice freezes on various substances cannot, however, be specified nearly so precisely.

According to de Bruyne (1951) the "first rule of adhesion" is the use of

"an adhesive that 'wets' the surface": therefore substances which can be wet by water can be expected to bond with ice. Bowden (1953) showed that plastics which do not wet easily have a low friction on ice at low temperatures and thus firmly established the experimental basis for an adhesion theory of ice friction: this constituted a tremendous advance in understanding the nature of the ice friction mechanism.

Bernal and Fowler (1933) have explained that the hydrogen bonds holding the oxygen atoms together in the ice structure are essentially electrostatic in character, which suggests a possibility that when ice freezes on to some substance the bond may be electrostatic. Weyl (1951) visualizes under the surface of the ice a gradual adjustment of the bonding for several hundred layers of molecules, to give a surface structure like that of water, and he points out that just as soon as the surface is covered, forces may be supplied at the surface which permit readjustment to the normal bonding of the interior of the ice. Wood absorbs water and supplies therefore the right type of bonding—presumably electrostatic. Teflon, on the other hand, which is hydrophobic and incidentally very resistant to chemical combination in general, apparently fails to supply the right type of forces to permit readjustment of the bonding below the surface of the ice, and so little adhesion occurs. Why bakelite is the only other organic material with a friction value comparable with that of teflon and why a material like polystyrene has a considerably higher friction value (Niven 1956) than either of these are questions dealing with fine points in the adhesion theory of friction which as yet remain unanswered. Since McConica (1950) found that the polarity of a ski wax was important in assessing its friction on snow, polarity may be a factor in explaining the behavior of the various plastics.

Metals, with their highly mobile valency electrons, are easily polarized: accordingly one can expect high friction between metal and ice at low temperatures unless the metallic surface has a protecting layer of oxide, difficult to polarize.

The adjustment of the bonding under the surface of the ice when covered by some substance having suitable forces, strictly speaking explains static friction more obviously than sliding friction, because, when a substance is sliding on ice, the molecules of the ice asperity are in a state of agitation and disarrangement before bonding occurs. But any substance which can form the bond by readjusting the forces under the surface will also form bonds with a resolidifying asperity in which the forces are in a state of disarrangement: therefore, one might expect that substances with high static friction on ice should also in general have high sliding friction.

The application of the adhesion theory of friction to rubber on ice presents a problem of special interest, because rubber is non-polar and should therefore have poor adhesion to ice: experiment shows, however, that its friction on ice is very high. Admittedly rubber requires special theoretical treatment (Denny 1953) to explain its friction on surfaces in general, on account of its exceptional elastic qualities. It has a high friction on dry surfaces: it can conform to the shape of the surface and thereby a weak adhesion bond can be multiplied more

often than is possible for a hard surface making contact only at the tips of asperities. Furthermore, because the rubber can shape itself over and around the ice asperity, it can, so to speak, grip any roughness—like the cogs on wheels interlocking. This grip may be somewhat like the grip which McConica (1950) describes as a "mechanical attachment" in discussing the behavior of plastic ski waxes on snow. However, no matter whether this grip on ice is a true adhesion or a mechanical attachment, it must let go if the asperity should melt. This explains some results obtained by the writer (Niven 1958) on rolling friction of rubber tires. At $-1°$ C the rolling friction was quite appreciable—indeed comparable with the friction at $-3°$ C—but at temperatures very close to 0° C it was small, indefinite, and independent of load at higher loadings.

CONCLUSION

While our knowledge of the adhesion bond between ice and the various substances is still imperfect, enough is known to indicate that adhesion plays a very important part in the ice friction mechanism. But adhesion is probably not the whole explanation because at temperatures just below 0° C the adhesion of the slider material to ice is of relatively small importance: both stainless steel and teflon slide excellently at these temperatures. It is at the colder temperatures that adhesion appears to become highly important: steel has then a far higher coefficient of friction than teflon. An idea somewhat like the old asperity theory of friction has been used in the above discussion to explain how the asperity is "prepared" by agitation and pressure for the adhesion bond to form by cooling. At temperatures only slightly below 0° C when conduction of heat from the asperity to the main body of the ice is very small, this adhesion or "refreezing" bond should be almost negligible and so the contribution to the frictional force made by one asperity pushing the other when they collide may quite reasonably be expected to become important.

Anything which helps to weaken the bonding in the ice so that the slider asperity can push forward after collision may be expected to lower friction, during the early stages of the mechanism: therefore, high temperature, high loading, and high impact (speed) should reduce friction. Anything which helps the bonding to occur between the slider asperity and the ice may be expected to increase friction, during the later stages of the mechanism: therefore, low temperature, low speed, and a slider substance which will wet or polarize easily should increase friction.

REFERENCES

BERNAL, J. D. and FOWLER, R. H. 1933. J. Chem. Phys. **1**, 515.
BOWDEN, F. P. 1950. Nature, **166**, 330.
——— 1953. Proc. Roy. Soc. A, **217**, 462.
BOWDEN, F. P. and FREITAG, E. H. 1955. Nature, **176**, 944.
BOWDEN, F. P. and HUGHES, T. P. 1939. Proc. Roy. Soc. A, **172**, 281.
BOWDEN, F. P. and RIDLER, K. E. W. 1936. Proc. Roy. Soc. A, **154**, 640.
BOWDEN, F. P. and TABOR, D. 1950. The friction and lubrication of solids (The Clarendon Press, Oxford).
DE BRUYNE, N. A. 1951. Structural adhesives (Lange, Maxwell and Springer Ltd., London), p. 3.

DENNY, D. F. 1953. Proc. Phys. Soc. B, **66**, 721.
KLEIN, G. J. 1950. National Research Council of Canada, Rept. MM-225.
MCCONICA, T. H. 1950. Sliding on ice and snow. Partial report to the Research and Development Branch Office of the Quartermaster General U.S. Army.
——— 1951. Wright Air Development Center, WADC Tech. Rept. No. 52-19, Appendix III.
NIVEN, C. D. 1954. Can. J. Phys. **32**, 782.
——— 1956. Can. J. Technol. **34**, 227.
——— 1956. Can. J. Phys. **34**, 262.
——— 1958. Can. J. Phys. **36**, 599.
PFALZNER, P. M. 1947. Can. J. Research, F, **25**, 192.
WAKEFIELD, W. E. 1938. Efficiency of logging sleighs for pulpwood operations in different types of terrain. Forest Products Laboratories, Dept. of Northern Affairs and National Resources, Ottawa. Project 107.
WEYL, W. A. 1951. J. Colloid Sci. **6**, 389.

ABSOLUTE STANDARDIZATION OF RADIOACTIVE NEUTRON SOURCES

I. ACTIVATION OF MANGANESE BATH[1]

K. W. GEIGER AND G. N. WHYTE[2]

ABSTRACT

The neutron emission rate of a Ra-α-Be source of the National Research Council has been measured by moderating the neutrons in manganese sulphate solution surrounding the source. Some of the thermal neutrons captured in the bath produce manganese-56, and the neutron emission rate is found from an absolute determination of the resulting activity. A number of corrections are discussed in detail. The value of $(3.22\pm0.05)\times10^6$ neutrons per second found for the neutron emission rate shows satisfactory agreement with the results of our intercomparison with the primary neutron standard of the National Bureau of Standards, and allows a tie-in with a series of international comparisons of primary neutron sources.

INTRODUCTION

For the determination of the emission rate, Q, of a radioactive neutron source three principal methods are available:

(1) The source neutrons can be thermalized in a moderator and counted absolutely; an integration over the volume of the moderator yields the total emission rate.

(2) Charged particles, associated with the neutron-producing reaction in the source, can be counted absolutely.

(3) The source and an absorber of known neutron absorption cross section can be introduced into a nuclear reactor and neutron emission and absorption balanced.

There are also a number of more indirect ways of determining the neutron emission rate of a source, such as by comparison with the neutron emission rate from an accelerator-induced reaction in which the associated particles can be counted absolutely. The $T^3(d, n)He^4$ reaction is commonly used for this purpose.

The present paper (I) describes a standardization using the first method, applied to the standard Ra-α-Be neutron source of the National Research Council (N.R.C.). A later paper (II) will deal with the standardization of a Po-α-F source by the second method.

PRINCIPLE OF THE MEASUREMENT

The measurement is based on the slowing down and capture of neutrons in a bath of aqueous manganese sulphate solution (O'Neal and Scharff-Goldhaber 1946; Alder and Huber 1949; DeJuren and Chin 1955*a*). If the

[1]Manuscript received October 24, 1958.
Contribution from the Division of Applied Physics, National Research Council, Ottawa, Canada.
Issued as N.R.C. No. 5063.
[2]Present address: Department of Physics, Queen's University, Kingston, Ontario.

bath is large enough, practically all the neutrons emitted by a source which is in the center are slowed down to thermal energies and captured, either by hydrogen or by manganese. From the ratio of the capture cross sections of manganese and hydrogen, one can calculate the fraction of the total captures that occur in manganese. Since manganese has only one stable isotope, Mn^{55}, each capture in manganese results in the formation of a nucleus of Mn^{56}, which decays with a half life of 2.58 hours. This can be readily detected through either its β- or γ-rays. If the neutron source is left in the bath until the manganese activity has built up to equilibrium, the number of manganese atoms disintegrating per second in the bath is equal to the number of neutrons captured per second by manganese nuclei, Q_{Mn}. If the total activity of Mn^{56} can be measured, the rate of neutron emission of the source is given by:

$$(1) \qquad Q = [\{(1+a)N_{Mn}\sigma_{Mn}+N_S\sigma_S+N_H\sigma_H\}/\{(1+a)N_{Mn}\sigma_{Mn}\}]Q_{Mn}$$

where N_{Mn}, N_S, and N_H are the numbers of manganese, sulphur, and hydrogen nuclei per cm^3, and σ_{Mn}, σ_S, and σ_H are the corresponding thermal-neutron capture cross sections. The factor $1+a$ arises because the manganese cross section does not vary exactly as $1/v$ in the epithermal region. A calculation following the method of DeJuren and Chin (1955*a*) shows that at the concentration used here the manganese captures 1.7% more neutrons than would a perfect $1/v$ absorber of the same thermal cross section. The factor $1+a$ thus amounts to 1.017.

When the solution is stirred, a mechanical integration of the induced activity over the whole volume of the bath is performed and the absolute activity of a known fraction of the solution yields the number of neutrons per second captured by manganese. However, the dimensions of the bath must be large to ensure that most of the neutrons are slowed down and captured. As a result most of the solution is exposed to only a very low thermal-neutron flux, and the average specific activity is too low to be measured directly. Instead a *relative* measurement of the activity is then made by inserting a radiation detector into the center of the bath. In order for the reading to be interpreted absolutely, the detector must be calibrated by observing its response when a known quantity of active manganese is mixed into the bath, so as to determine the number of disintegrations per second of Mn^{56} in the bath corresponding to 1 count per minute in the immersed detector.

APPARATUS AND EXPERIMENTAL PROCEDURE

The manganese sulphate bath was contained in an open cylindrical steel tank, 85 cm in diameter, coated with neoprene to prevent corrosion. The tank was filled to a depth of 85 cm with a nearly saturated aqueous solution of manganese sulphate containing 30.31% $MnSO_4$ by weight. The exact concentration was determined from the density of the solution and its tabulated dependence on concentration (International Critical Tables 1928).

The source to be calibrated was the N.R.C. radium-α-beryllium source No. N-200-1, containing 183 mg of radium element in the form of the carbonate

and 2000 mg of beryllium powder, pressed into a pellet to ensure physical stability. The source was prepared in 1948. The pellet is contained in a cylindrical monel capsule of 2 mm wall thickness, both the diameter and the height being 19 mm.

In order to make a measurement the source was enclosed in a thin-walled polyethylene container and suspended in the center of the solution for a period of several hours. After removal of the source, the bath was stirred for 15 minutes by means of a 1/8 horsepower electric mixer.

The relative activity of the bath was measured by inserting a commercial "multicellular" Geiger–Müller tube, protected by a brass shell, into the center of the solution. This counter consisted of several bismuth-cathode tubes in parallel, enclosed in an outer envelope 30 cm long by 5 cm in diameter, which combines high γ-efficiency with short dead time. The response of the counter was checked periodically with a small radium source and found to be stable to within better than $\pm 0.3\%$.

Five or more 10-minute counts were taken and the rates corrected for decay of Mn^{56} back to the time of source removal. The published half life (DeJuren and Chin 1955*a*) of 2.58 hours was used, and this fitted the data within the statistical uncertainties. Two runs over several half lives also confirmed this figure and showed that no other radioactive γ-emitter was produced in the bath. Since irradiation with the source took place for only approximately 15 hours, a further small correction had to be made to obtain the counting rate for irradiation to equilibrium.

THE SZILARD–CHALMERS PROCESS AND ABSOLUTE STANDARDIZATION OF Mn^{56}

To calibrate the counter in the bath a known amount of Mn^{56} has to be added to the inactive bath. The best method to determine the absolute Mn^{56} activity is by 4π β-counting; however, the specific activity of the Mn^{56} solution has to be high to make self-absorption of the source deposit in the 4π counter negligible. This was achieved by using the Szilard–Chalmers (1934) process.

Approximately 600 g of $Ca(MnO_4)_2$ in 900 ml H_2O were irradiated overnight with a 1 g Ra-α-Be neutron source. The activated manganese leaves the permanganate ion by recoil due to neutron capture γ-rays and forms insoluble MnO_2, which was separated by filtering through a glass filter, medium grain. The active MnO_2 residue was dissolved in 8 ml of 6 *M* HCl. A few drops of H_2SO_4 were added and the solution evaporated to dryness until no more SO_3 vapors escaped. The few milligrams of resulting $MnSO_4$ were dissolved in H_2O, filtered, and made up to 50 ml. This represented the stock solution. The Szilard–Chalmers process yielded about 60% of the available active manganese and an enrichment in Mn^{56} isotope by a factor of 5000.

Three samples of 0.05 ml were pipetted for 4π counting. The samples were prepared following the method of Pate and Yaffe (1955) on gold-plated VYNS foil (approximately 5 $\mu g/cm^2$ foil and 15 $\mu g/cm^2$ Au) using insulin as wetting agent. Drying was carried out in an NH_3 atmosphere under infrared light. This produced a very finely grained MnO_2 deposit which was about 100 $\mu g/cm^2$ thick.

The 4π counter was operated in the proportional region under continuous CH_4 gas flow. Counting plateaux with no statistically significant variation could be obtained over a range of 200 volts by using a non-overloading pulse amplifier of 0.2-mv input sensitivity. Since the dead time of the amplifier was 5 μsecond, resolving time corrections were negligible for the maximum counting rates (8000 c.p.m.) which occurred.

Absorption in the source mount and in the active deposit should be negligible at the thicknesses used (Baker and Katz 1953; DeJuren and Chin 1955*a*). It is estimated that the uncertainty of the standardization of Mn^{56} is $\pm 1\%$, mainly due to pipetting errors and non-uniformity of the source deposit.

The 4π counting yielded the activity per milliliter in the stock solution. The rest of the solution was stirred into the $MnSO_4$ bath, and the response of the multicellular counter to the known Mn^{56} activity was observed. This calibrated the counter in the bath. Such calibration of the counter–bath assembly was made before each activation of the bath by the neutron source to be standardized; the actual value was around 66 Mn^{56} disintegrations per second for one count per minute as recorded by the multicellular counter. This figure changed very little for the various runs. The neutron emission rate from the source was then calculated using Equation 1 with the cross-section values as listed in Table I. With a $MnSO_4$ concentration of 30.31% the capture rate becomes

$$Q = 1.985\, Q_{Mn}.$$

Since nearly half of the total neutrons are being captured in manganese the contribution of the cross-section error in the final result is reduced by a factor of two.

TABLE I

Thermal neutron absorption cross sections

	Cross section, barns	Reference
H^1	0.332 ± 0.002	Hughes and Schwartz 1957
S	0.49	Hughes and Harvey 1955
Mn^{55}	$13.2_5 \pm 0.2$ (Best estimated value)	Hughes and Harvey 1955 DeJuren and Chin 1955*b* Cummins and Spurway 1957

CORRECTIONS

If no neutrons were lost by other processes, the capture rate, Q, in the bath could be equated to the emission rate of the source. Actually small losses of several kinds occur, and these must first be taken into account.

Fast-Neutron Capture

At energies above a few million electron volts, neutrons can undergo nuclear reactions with the oxygen and sulphur present in the bath. The neutrons that interact in this way do not reach thermal energies and hence do not contribute to the activation of the manganese. This loss of fast neutrons must be estimated and included in the total emission rate of the source.

The fast-neutron loss in water for the case of a Ra-α-Be source has been determined experimentally by de Troyer and Tavernier (1954), who compared the slowing-down density of 1.44-ev neutrons in water and paraffin oil (which contains no oxygen). Their results indicate that $(2.2\pm0.3)\%$ of the neutrons emitted by the source were absorbed at high energies by the oxygen in the water.[3] In the absence of similar measurements in $MnSO_4$, the additional absorption in the sulphur can only be estimated from a comparison of the sulphur- and oxygen-absorption cross sections over the limited range of energies for which they have been measured. If the resonance escape probability in water is put equal to 0.978 on the basis of the measurements of de Troyer and Tavernier, an average absorption cross section in oxygen between 4 Mev (below which it is known to be negligible) and 13 Mev (the upper limit of the Ra-α-Be neutron spectrum) can be deduced; the result is 0.2_9 barns. The (n, p) and (n, α) cross sections in sulphur have been measured between 2 and 4 Mev (Hughes and Harvey 1955); their sum increases from 0 to about 0.45 barns in this region. If, therefore, the average cross section between 2 and 13 Mev is assumed to be between $\frac{1}{2}$ and 4 times the oxygen cross section, the total fast neutron absorption in the $MnSO_4$ solution amounts to $(2.8 \pm0.4)\%$.

Absorption in the Neutron Source

A correction must also be made for the absorption of thermal neutrons in the source itself. De Troyer and Tavernier (1954) have determined the amount of this absorption by measuring the effect of additional absorbing material placed around the source on the neutron density integrated over the volume of the bath. For the case of a source with a total absorption cross section of 1.0 cm^2 in a bath of pure water they found the absorption to be $(0.75\pm0.3)\%$ of the emitted neutrons. The effect of the manganese in the bath used here is to reduce the thermal-neutron flux at any point by a factor of roughly

$$\Sigma_{H_2O}/(\Sigma_{H_2O}+\Sigma_{Mn}) = 0.495$$

from the value in pure water. (The Σ's are the macroscopic absorption cross sections.) The total absorption cross section of the present source was 1.16 cm^2, so that the absorption in this case would amount to

$$0.75\times0.495\times(1.16/1.0) = (0.4\pm0.2)\%.$$

Escape

A third way in which neutrons can be lost is by escape through the boundaries of the bath. The magnitude of this loss has been estimated in two ways:

(1) Larsson (1954) has found that the thermal-neutron density around a source in a water bath, multiplied by the square of the distance from the source, falls off exponentially at large distances. On the assumption that the area under the portion of the exponential curve extending beyond the

[3]De Troyer and Tavernier actually obtained a figure of 2.5%, but DeJuren (1958) pointed out a small error in their derivation of this figure.

boundaries of the tank represents the number of neutrons escaping from the tank, it is possible to use this relationship to estimate the fraction of the neutrons escaping. On the basis of Larsson's data this fraction is given by

$$f = 1.5\, e^{-0.10r}$$

where r is the distance from the source to the edge of the tank in centimeters. Taking an average of this expression over the boundaries of the tank used in the present experiment gives $f = (1.4 \pm 0.4)\%$. Since Larsson found that the addition of a neutron absorber (H_3BO_3) to the water in the tank did not alter this fraction appreciably, the above figure should also apply to the $MnSO_4$ solution. (The slowing-down lengths in water and in the $MnSO_4$ solution do not differ by enough to affect this result.)

(2) Rumsey and Ward (1948) have measured the total neutron capture in $MnSO_4$ baths of various sizes. Their data can be fitted by the following expression for the fractional escape:

$$f = 1.9\, e^{-0.105r}.$$

This also yields a correction of 1.4% in the present case.

RESULTS AND DISCUSSION

The results of seven runs, each consisting of a calibration of the counter in terms of disintegrations per second in the bath, followed by a measurement of the activity induced by the standard source, are as follows:

Date	(Neutrons/sec)
7/5/58	3.22×10^6
14/5/58	3.21
23/5/58	3.21
16/7/58	3.22
23/7/58	3.19
30/7/58	3.26
8/8/58	3.23
	Mean 3.22×10^6

The following are the estimated errors in the final result arising from various sources:

Random counting errors	0.3%	Correction for absorption in neutron source	0.2%
Systematic errors in 4π counting	1.0	Escape correction	0.4
Manganese concentration in bath	0.3	Resonance absorption in manganese	0.2
Capture cross sections	0.8		
Correction for fast-neutron capture	0.4		

The result for the emission rate of the source is

$$Q = (3.22 \pm 0.05) \times 10^6 \text{ neutrons/second for July 1958.}$$

This can be compared with a value obtained by Chin and Caswell (1956), who compared the Canadian source with the National Bureau of Standards primary Ra-Be photoneutron standard (NBS I) in January 1956 using a manganese sulphate bath. After correction for fast-neutron capture in the case of the (α, n)-source a neutron emission rate of $(3.13 \pm 0.09) \times 10^6$ per second was obtained for the Canadian source. Corrected for growth of Po^{210}

its emission rate would be 3.16×10^6 neutrons per second in July 1958. This is 2% lower than the present value, but still within the uncertainties of the measurements. With this comparison the N.R.C. source is tied into an international comparison of primary standards which was carried out by Larsson (1958). It is planned to have the Canadian standard compared directly with a number of other national neutron standards, most of which are of the Ra-α-Be type.

REFERENCES

ALDER, F. and HUBER, P. 1949. Helv. Phys. Acta, **22**, 368.
BAKER, R. G. and KATZ, L. 1953. Nucleonics, **11** (2), 14.
CHIN, J. and CASWELL, R. S. 1956. Unpublished.
CUMMINS, J. D. and SPURWAY, A. H. 1957. United Kingdom Atomic Energy Authority Report A.E.R.E. R/M 100.
DEJUREN, J. 1958. Unpublished.
DEJUREN, J. A. and CHIN, J. 1955*a*. J. Research Natl. Bur. Standards, **55**, 311.
——— 1955*b*. Phys. Rev. **99**, 191.
HUGHES, D. J. and HARVEY, J. A. 1955. Brookhaven Natl. Lab. Report BNL-325.
HUGHES, D. J. and SCHWARTZ, R. B. 1957. BNL-325, Suppl. 1.
INTERNATIONAL CRITICAL TABLES. 1928. Vol. 3 (McGraw-Hill Book Co., New York), p. 68.
LARSSON, K.-E. 1954. Arkiv Fysik, **7**, 323.
——— 1958. J. Nuclear Energy, **6**, 322.
O'NEAL, R. D. and SCHARFF-GOLDHABER, G. 1946. Phys. Rev. **69**, 368.
PATE, B. D. and YAFFE, L. 1955. Can. J. Chem. **33**, 610.
RUMSEY, V. H. and WARD, A. G. 1948. Unpublished. *As quoted by* Hanna, G. C. and Runnalls, O. J. C. 1956. Can. J. Phys. **34**, 959.
SZILARD, L. and CHALMERS, T. A. 1934. Nature, **134**, 462.
DE TROYER, A. and TAVERNIER, G. C. 1954. Acad. roy. Belg., Bull. classe sci. **40**, 150.

FOUR-DIMENSIONAL ELECTROMAGNETIC RADIATORS[1]

H. E. SHANKS AND R. W. BICKMORE

ABSTRACT

This paper describes an entirely new and fundamental technique which has wide application to many fields of the antenna art as well as important system implications. The essence of this philosophy is the utilization of the time domain as an additional variable with which to control antenna radiation characteristics. One way to accomplish this is by periodically time-modulating one or more of the antenna parameters (aperture excitation, aperture shapes, frequency, phase distribution, aperture size, etc.) in a prescribed way. The result of this modulation is a radiation pattern whose characteristics are periodically changing as a function of time. By virtue of the periodic nature of the pattern fluctuations, an infinite number of independent information channels are available corresponding to the harmonic frequency components of the modulated patterns. Since each harmonic will, in general, have a different space factor associated with it, proper data processing can be utilized to provide simultaneous operation of a single antenna in a variety of modes. In addition, this technique allows synthesis of antenna radiation characteristics which might be unobtainable by conventional methods.

INTRODUCTION

During the past 10 or so years the development of radiating systems has advanced to the point where a multitude of antenna techniques are available for any particular operational requirement. These techniques embrace the use of one-, two-, and even three-dimensional source distributions. However, even with all three spatial co-ordinates available, increasingly stringent system specifications cannot always be met. In other words, an additional degree of freedom—a fourth dimension—is needed to realize an order of magnitude improvement in antenna operation. The fourth dimension with which we are concerned here is just that so often mentioned in the popular literature—time.

The essence of four-dimensional radiating systems is the periodic time modulation of one or more antenna parameters so as to increase the system information handling capacity or realize advanced radiation pattern characteristics. The modulated parameter can be the physical dimensions of the antenna, the energy distribution function over the antenna, the frequency or even the physical location of the phase center. An antenna technique investigated in the past (Cutrona 1955) is a practical example of the latter form of modulation. In this situation the antenna is physically moved and the received energy is processed to attain improved pattern characteristics. The necessity for processing is a function of the extremely low modulation rates (one or two cycles per second) involved. If the antenna modulation occurs at a much higher rate, processing takes place in a spatial sense and thus simplifies the system operation greatly. Application of these concepts to other parameters mentioned above results in exploitation of other pattern characteristics, for example, direction of beam collimation or sidelobe level.

[1]Manuscript received October 27, 1958.
Contribution from the Microwave Laboratory, Hughes Aircraft Company, Culver City, California.

GENERAL REMARKS

In its most general form an antenna may be described as a transducer between a spatial distribution of energy and a temporal distribution of information. Although this definition is largely academic, it will serve to point up the differences between three-dimensional and four-dimensional operation. Under this definition an antenna may be considered as a two-terminal pair network with normal operation restricted to stationary and linear transfer functions. However, with time as an additional degree of freedom, the equivalent network is no longer stationary and may in fact be non-linear, the output information thus being a function of the non-stationary (modulated) nature of the antenna as well as the input signal.

If a wave of energy is incident on an antenna whose parameters are modulated in a periodic manner, the voltage across the output terminals will be of the following form,

$$E(\theta, t) = A\, g(t)\{b_0(\theta)+b_1(\theta)\cos\omega_A t+b_2(\theta)\cos 2\omega_A t+\ldots\}e^{i\omega t} \tag{1}$$

where A contains the radial dependence, θ denotes the spatial variation in the signal, and the Fourier series containing the $b_n(\theta)$ is the time-dependent radiation pattern with ω_A the fundamental modulation frequency. In (1), $g(t)$ represents the input information, which takes different forms depending on the operational use; for radar $g(t)$ is of a pulsed nature and contains the range information while in a communication link $g(t)$ might represent an audio voice spectrum. In any case the basic time-modulated character of the pattern information is identical.

For general modulation functions the space factors, $b_n(\theta)$ associated with the harmonically related frequency components $n\omega_A$ will be different. Because of the orthogonal nature of the Fourier terms each term may be considered as an independent information channel capable of carrying different spatial dependent information. The implication of this is clear; if an antenna is properly modulated, a multitude of spatial patterns is generated, each one associated with its characteristic frequency component. By proper processing of this complex signal, the individual patterns may be separated and each one utilized in a normal manner. It is thus obvious that the information capacity of a single antenna can be enormously increased by the use of time domain techniques.

The preceding remarks present a qualitative picture of four-dimensional antenna operation. However, the question arises: How must the antenna be modulated to realize the proper spatial pattern $b_n(\theta)$? Consider a distribution of radiating sources spread over a surface S_0. The radiation pattern is

$$g(\theta) = \int_{S_0} \Psi(s_0)\, G(\theta|s_0)\, ds_0 \tag{2}$$

where $\Psi(s_0)$ is the complex distribution of energy over the surface, $G(\theta|s_0)$ is the Green's function characteristic to the surface S_0 and θ is the space variable. If $\Psi(s_0)$ is allowed to vary with time, that is $\Psi(s_0, t)$, then we desire the form of this variation to realize a prescribed set of $b_n(\theta)$'s in the radiation pattern $g(\theta, t)$.

Under these conditions (2) may be rewritten with the time variation included as

$$g(\theta, t) = \int_{S_0} \Psi(s_0, t)\, G(\theta|s_0)\, ds_0 \tag{3}$$

and if the variation is periodic, with frequency ω_A, the distribution may be expanded

$$\Psi(s_0, t) = \phi_0(s_0) + \phi_1(s_0) \cos \omega_A t + \phi_2(s_0) \cos 2\omega_A t + \ldots \tag{4}$$

Substituting (4) into (3) results in

$$g(\theta, t) = \sum_{n=0}^{\infty} \left\{ \int_{S_0} \phi_n(s_0)\, G(\theta|s_0)\, ds_0 \right\} \cos n\omega_A t \tag{5}$$

and if this expression is compared term by term with the time modulated pattern of (1), there obtains

$$b_n(\theta) = \int_{S_0} \phi_n(s_0)\, G(\theta|s_0)\, ds_0. \tag{6}$$

In other words, each harmonic coefficient $\phi_n(s_0)$ in the distribution gives rise to the corresponding coefficient $b_n(\theta)$ in the pattern, so that given the desired $b_n(\theta)$'s the distribution $\Psi(s_0, t)$ is obtained from (6) and (4).

The foregoing analysis and discussion presents the basic ideas of time domain operation. Nothing has yet been said about the data processing schemes to realize these objectives. Also, only modulation of the energy distribution parameters has been considered. These topics will be discussed in later sections.

APPLICATIONS

The operation of an antenna system with four-dimensional control will undoubtedly find numerous applications to military as well as commercial operations. Some of the most interesting characteristics which may be achieved by these techniques are: multipattern operation, simultaneous scanning, and sidelobe suppression. These applications will be discussed in some detail below.

Multipattern Operation

In many instances it would be highly desirable to have a single antenna capable of operation in a variety of widely different pattern modes. For example, in a ground-based radar, broad search capabilities might be desired along with accurate detection and tracking capacity. This capability is readily realized through time domain techniques by modulating the antenna in such a way as to cause the space factors associated with two harmonic pattern frequencies to be a pencil beam and a $\csc^2$ or fan type pattern respectively. These pattern features would exist simultaneously and, by proper processing of the complex signal, either mode or both modes of operation can be selected.

As a simple example of this concept, consider a linear array of length $2l_0$ from which a pencil beam and a $\csc^2$ pattern are desired. Then, in keeping with equation (1) and considering only the time-varying pattern, let

$$b_0(\theta) = \frac{A_0 \sin (k\, l_0 \sin \theta)}{\sin \theta}$$

and

$$b_1(\theta) = A_1 \csc^2 \theta.$$

The complete time-varying pattern is then

$$g(\theta, t) = \frac{A_0 \sin (k\, l_0 \sin \theta)}{\sin \theta} + A_1 \csc^2 \theta \cos \omega_A t \tag{7}$$

and the antenna modulation necessary to realize this characteristic can be found by the procedure leading to equation (6). Although the pattern features shown in (7) are to be realized from a single array, it is enlightening to consider this array as broken down into two component elements. Figure 1 illustrates two linear arrays in symbolic form. One of these arrays is unmodulated and

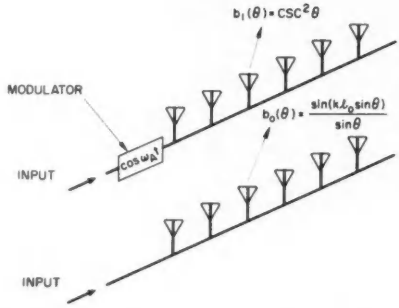


FIG. 1. Symbolic representation of component arrays for multipattern operation.

has the distribution necessary to give the first term of (7) (a pencil beam). The input to the second array is modulated as $\cos \omega_A t$ with a distribution to achieve the $\csc^2 \theta$ characteristic. The pattern resulting from the combination is just that given by equation (7). These two arrays are then superimposed into an equivalent single array and the modulated distribution is achieved by a means which is, for the present, unspecified. It should be readily seen that similar reasoning could be applied to planar array antennas as well as dish-reflector type systems with similar results.

Simultaneous Scanning

In many military applications it would be extremely useful to have a non-scanning antenna which would detect and locate targets over a wide angular region. This is again readily available using time domain operation; in this case, the space factors $b_n(\theta)$ being pencil beams each pointed in a slightly different direction. Again considering a linear array of length $2l_0$, let

$$b_n(\theta) = A_n \frac{\sin \{k\, l_0 [\sin \theta - \sin \theta_n]\}}{[\sin \theta - \sin \theta_n]}, \tag{8}$$

that is, the pointing direction for each harmonic space factor different and directed at θ_n.

The time-varying pattern under this condition is

$$g(\theta, t) = \sum_{n=0}^{N} A_n \frac{\sin \{k\, l_0 [\sin \theta - \sin \theta_n]\}}{[\sin \theta - \sin \theta_n]} \cos n\omega_A t \tag{9}$$

where the A_n are adjusted to make the return from the peak of each space pattern be equal. The number N and the values θ_n would be chosen consistent with the maximum coverage desired and the necessary discrete angular accuracy. Data processing would separate the harmonics and feed the angular information to corresponding channels which would be used to drive the display apparatus.

Sidelobe Suppression

Preceding applications of the time domain have utilized modulation of the energy distribution function and have considered the use of more than one frequency harmonic. For the present application we consider only the time average pattern, $b_0(\theta)$, and determine modulation techniques which result in the time average having improved pattern characteristics (in particular low sidelobes). The utility of this concept is that the extreme mechanical tolerances necessary to realize low sidelobes in a non-modulated antenna may be relaxed by virtue of the additional degree of freedom provided by the time variation.

This philosophy may be understood by considering a simple linear array as illustrated in Fig. 2. If this array has a uniform distribution, the sidelobe level

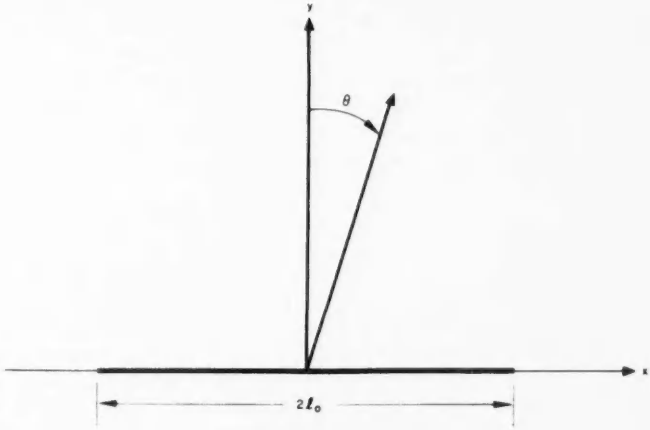


FIG. 2. Co-ordinate system for linear array.

is 13.2 db and it will be shown that considerable reduction in this figure is obtained if the array length is modulated in a square wave manner.

The pattern for the linear array of Fig. 2 is

$$g(\theta) = A \frac{\sin (k\, l_0 \sin \theta)}{\sin \theta} \tag{10}$$

and has a sidelobe level of 13.2 db. If it is assumed that the half array length is modulated in time as

$$l = l_0[1+\beta f(t)] \tag{11}$$

where $f(t)$ is the time modulation and β is a length parameter, the time-varying pattern is

$$g(\theta, t) = A \frac{\sin \{k\, l_0\, [1+\beta f(t)] \sin \theta\}}{\sin \theta}. \tag{12}$$

The time average is obtained by the common integral expression

$$\langle g(\theta, t)\rangle = b_0(\theta) = \frac{1}{T}\int_0^T g(\theta, t)\, dt \tag{13}$$

where T is the modulation period. When (12) is substituted into (13) and the integral evaluated with $f(t)$ a square wave

$$f(t) = \begin{cases} 1 & 0 \leqslant t \leqslant T/2 \\ -1 & T/2 \leqslant t \leqslant T \end{cases}$$

the resulting time average pattern is

$$b_0(\theta) = A \frac{\sin \{k\, l_0 \sin \theta\}}{\sin \theta} \cos (k\, l_0\, \beta \sin \theta). \tag{14}$$

The multiplicative term $\cos (kl_0\, \beta \sin \theta)$ has the effect of an additional pattern taper resulting in lowered sidelobes. The parameter β is the controlling quantity in setting the sidelobe level. Figure 3 displays a power plot of equation (14) for judicious choices of this parameter. The optimum condition is seen to correspond to a β value of 0.25 giving a sidelobe level of approximately 23 db—a 10-db reduction due to time modulation. It should be emphasized that this order of magnitude reduction in sidelobe level is achieved by switching only between two differing patterns and with perhaps the simplest form of modulating function.

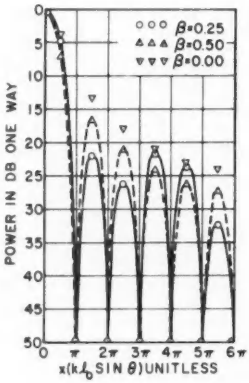


FIG. 3. Patterns for array with time-varying length.

This analysis has definitely established the sidelobe suppression capability of time domain techniques; however, the particular example realizes only a 10-db reduction. It is desirable to obtain a synthesis procedure by which a given sidelobe level can be realized. Considering again a uniform linear array, the integral form of the time average pattern can be expressed as

$$\langle g(\theta, t)\rangle = \frac{1}{T}\int_0^T \int_{-l(t)}^{l(t)} e^{jkx \sin\theta}\, dx\, dt \tag{15}$$

and we wish to determine the length variation $l(t)$ so as to realize a desired pattern form, $\langle g(\theta, t)\rangle$, with given sidelobe levels.

Such a procedure has been worked out for the linear array, and applied to obtaining a pattern with controllable sidelobes. A simple pattern of this type has been given in the literature (Taylor 1952). This pattern is

$$g(u) = \frac{\sin\left(\pi\sqrt{u^2-B^2}\right)}{\left(\pi\sqrt{u^2-B^2}\right)} \tag{16}$$

where u is the normalized space variable and B is the sidelobe controlling factor; the first sidelobe voltage ratio being given by

$$4.6 \sinh \pi B/\pi B.$$

The synthesis procedure is based on the analogy of (15) and the expression for the principal plane radiation pattern of a two-dimensional aperture. This expression is,

$$g(\theta) = \int_0^h \int_{-f(y)}^{f(y)} \left\{ e^{jkx \sin\theta} dx \right\} dy \tag{17}$$

where $f(y)$ represents the shape of the aperture and h the maximum y dimension. Comparison of (15) and (17) indicates the equivalence of t and y through the relations

$$f(y) \sim l(t), \qquad h \sim T. \tag{18}$$

Thus, it is clear that time modulation of the length of a linear array is equivalent to adding a second dimension of control. Having established the equivalence of (18) it is possible to use two-dimensional aperture synthesis methods to arrive at the function $f(y)$, (and hence $l(t)$ by (18)) to realize the pattern of (16). Figure 4 shows the general form of $l(t)$ plotted on a universal time scale. The parameter which controls the sidelobe level in this case (in analogy with B of (16)) is the ratio of the period T to the pedestal height denoted by unity in the figure.

SYSTEM REQUIREMENTS

The previous section has described a number of applications of the time domain to advanced radiating system characteristics. These techniques have all been based on the ability of a time-modulated antenna to either increase the information-handling capability of the system or to realize improved

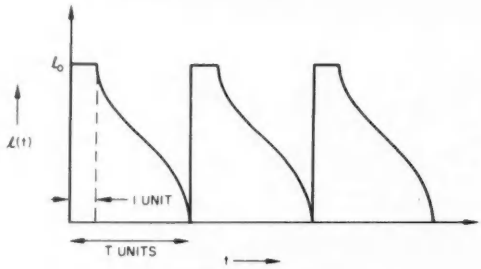


FIG. 4. Length variation for Taylor pattern.

radiation pattern characteristics. In order to capitalize on the increased versatility provided by these antenna techniques, due consideration must be given to the accompanying system and detection requirements. What modification to present system design procedures are necessary to accommodate four-dimensional antennas? How must the modulation frequency be related to other signal frequencies? All of these questions are of basic importance in the application of time domain antennas to any practical radiating system.

It should be obvious that the solution to these and other pertinent problems could constitute a complete paper in themselves. Also, the particular operational use (radar, communication, radio astronomy, etc.) under consideration is intimately connected with these solutions. As a result it will not be possible to cover all the system aspects in detail; instead a few general remarks will be given to illustrate possible system solutions. In addition, a detailed analysis of four-dimensional antenna operation in a voice communication link will be carried out.

Basically the problem presented to the detection system is one of separating a large number of harmonic frequency components and at the same time retaining the input signal information. This statement again points up the differences which arise because of the various operational uses, the input signal nature being different in each case. However, it seems reasonable to assume that a normal receiver and detector in conjunction with a bank of appropriate filters would be a starting point for any application of these

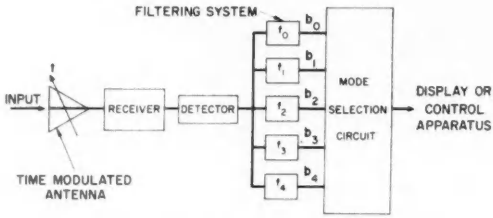


FIG. 5. Block diagram of system for four-dimensional antenna.

techniques. Figure 5 shows a symbolic block diagram of such a system. The incoming signal is incident on the antenna, modulated in the proper fashion, and then fed to a conventional receiver. From the receiver the information is sent to the filters for separation and then to the mode selection circuit which feeds the desired combination of signals to the display or control apparatus. The harmonic frequency separations and the filter bandpass characteristics would be chosen consistent with the signal information to be processed. Alternative schemes utilizing coherent detection techniques might also be called for in particular applications. It should be re-emphasized that this system is not proposed as a solution to a particular operational requirement but only as a symbolic illustration of potential system concepts.

A more detailed analysis of a particular system will prove enlightening. Consider a voice communication antenna in which sidelobe suppression is artificially realized by modulation techniques. The input angular carrier frequency is chosen as ω_0, the voice spectrum extends from 100 c.p.s. to 3000 c.p.s., and the fundamental modulation frequency is ω_A. Assuming a superheterodyne receiver, the carrier frequency will be reduced to an intermediate frequency which will be denoted by ω_{IF}. Under these conditions the voltage incident on the detector is given by

$$E(\theta, t) = Ag(t)\{b_0(\theta)+b_1(\theta)\cos\omega_A t+b_2(\theta)\cos 2\omega_A t+\ldots\}\cos\omega_{\mathrm{IF}} t. \tag{19}$$

Out of this complex signal we desire only the sidelobe suppressed pattern, along with the voice signal represented by $g(t)$. Mathematically that is,

$$Ag(t)\, b_0(\theta)\cos\omega_{\mathrm{IF}} t. \tag{20}$$

This is essentially a problem of designing the bandwidth of the receiver to pass the information of (20) while rejecting all remaining terms of (19). Rather than determine the necessary bandwidth, let us fix the bandwidth just sufficient to pass the upper voice frequency and determine the lower limit on ω_A to achieve rejection. Under this condition the frequencies corresponding to the upper limit of the voice spectrum are $\omega_{\mathrm{IF}}\pm 3000$ c.p.s. for dual sideband operation and the lower limit of the error frequencies (ω_A) is $(\omega_{\mathrm{IF}} \pm\omega_A)\mp 3000$ c.p.s. We desire the lowest error frequency to be outside the upper voice frequency; mathematically,

$$\omega_{\mathrm{IF}}\pm\omega_A\mp 3000 \text{ c.p.s.} \geqslant \omega_{\mathrm{IF}}\pm 3000 \text{ c.p.s.} \tag{21}$$

or

$$|\omega_A| \geqslant 6000 \text{ c.p.s.} \tag{22}$$

In other words for a receiver bandwidth of 3000 c.p.s. (SSB) or 6000 c.p.s. (double sideband) the lowest possible antenna modulation frequency is 6000 c.p.s. Thus time domain patterns in this particular case can be detected with no modification in the detection system with due regard to the modulation frequency ω_A. Figure 6 is a graphical illustration of the situation just analyzed above.

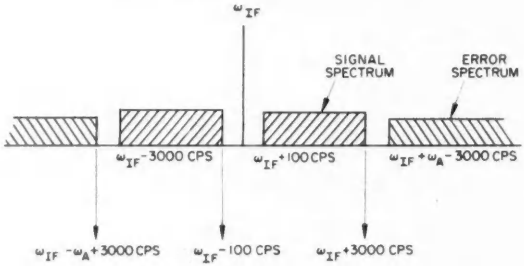


FIG. 6. Frequency spectrum for time-modulated communications antenna.

PHYSICAL CONFIGURATIONS AND EXPERIMENTAL RESULTS

Many of the important aspects of four-dimensional radiating systems have been discussed earlier; there only remains to suggest possible physical antenna configurations capable of realizing the time modulation characteristics which might be required in various applications. These configurations will, obviously, take quite different forms depending on the type of antenna under consideration (diffracting, reflecting, or refracting). Array type antennas are a particularly attractive structure because of the capability of point by point control over the energy distribution function. This feature is very appropriate when modulation characteristics are desired as in the present instance. Reflecting and refracting type systems are not as versatile as the array antenna, but still lend themselves to modulation capacity. In the case of the reflector, the primary feed characteristics may be modulated in time while for the refracting structure, the lens material or position may be varied. Thus, modulation parameters are available in almost any type of antenna system.

One of the simplest embodiments of the four-dimensional concept is found in the square wave switched linear array. A center-fed slot array is depicted in Fig. 7. The end portions of this array are fed through ferrite switches which

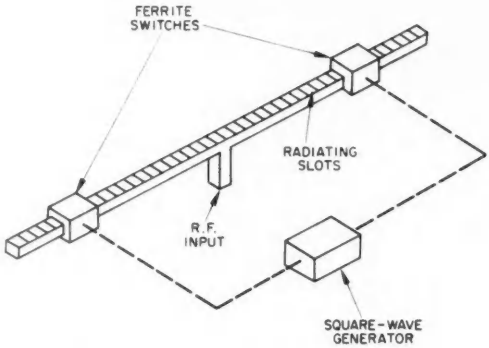


FIG. 7. Square-wave switched linear array.

are in turn driven by a square wave generator, resulting in an effective variation in the array length. A system of these arrays could be readily stacked and synchronously modulated to form a two-dimensional radiating surface with improved sidelobe characteristics. As has been shown earlier, sidelobe suppression greater than 10 db can be achieved from linear arrays by considering modulations other than square wave.

The most general and versatile four-dimensional system is one in which each element is individually controlled in amplitude and phase by its own modulating element. This is exhibited symbolically in Fig. 8 showing a two-dimensional array of radiators each fed through a single modulating element.

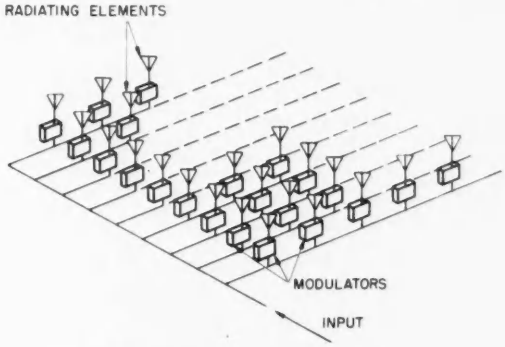


FIG. 8. Generalized four-dimensional antenna.

For low-frequency applications dipole elements would probably be used with ferrite or traveling wave tube modulating devices. At higher frequencies slot array techniques would be more appropriate. In this case each slot could be fed through its own branch line containing the modulation element or very recently developed techniques using ferrite excited slots (Shanks 1958) might be utilized. This concept of a large number of individually controlled elements represents the most refined four-dimensional structure and is capable of realizing any of the pattern characteristics mentioned in earlier sections.

In order to illustrate the range of applicability of time domain operation, experimental tests have been made on a readily available modulated antenna. This antenna is just the common parabolic dish such as is used in airborne fire control systems. The modulation is provided by a nutating primary horn used in normal operation to achieve conical lobing. Intuitive considerations indicated that if the horn were nutating, the time average pattern (d-c.) would display a pencil beam shape while the spatial pattern associated with the fundamental frequency component would be characterized by a null in the on axis position. Figures 9–11 show the experimental patterns and bear out this reasoning. Figure 9 is a reference pattern taken with the horn stationary while Fig. 10 is the time average pattern with the feed rotating. The split beam pattern shown in Fig. 11 is the pattern obtained at the fundamental

modulation frequency and is seen to resemble the difference pattern of a monopulse antenna. In all cases the horn was rotating at 133 c.p.s. The actual experimental antenna is displayed in Fig. 12 along with the tunable bolometer amplifier used to filter the proper frequency components. These results are an

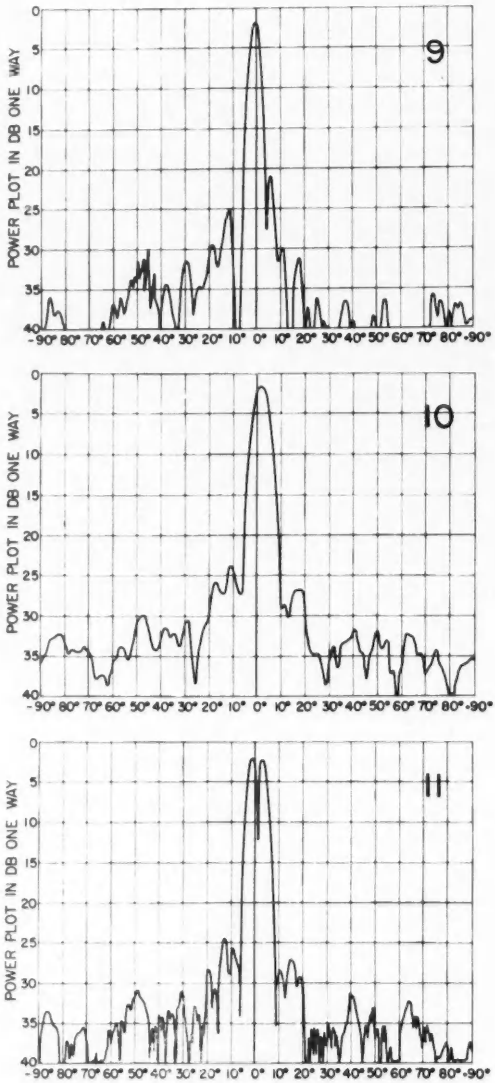


FIG. 9. Reference pattern of conical-lobed antenna.
FIG. 10. Direct-current response of conical-lobed antenna.
FIG. 11. The 133-c.p.s. response of conical-lobed antenna.

excellent example of multipattern operation described earlier, a sum and difference type pattern both being generated simultaneously and associated with different frequency components.

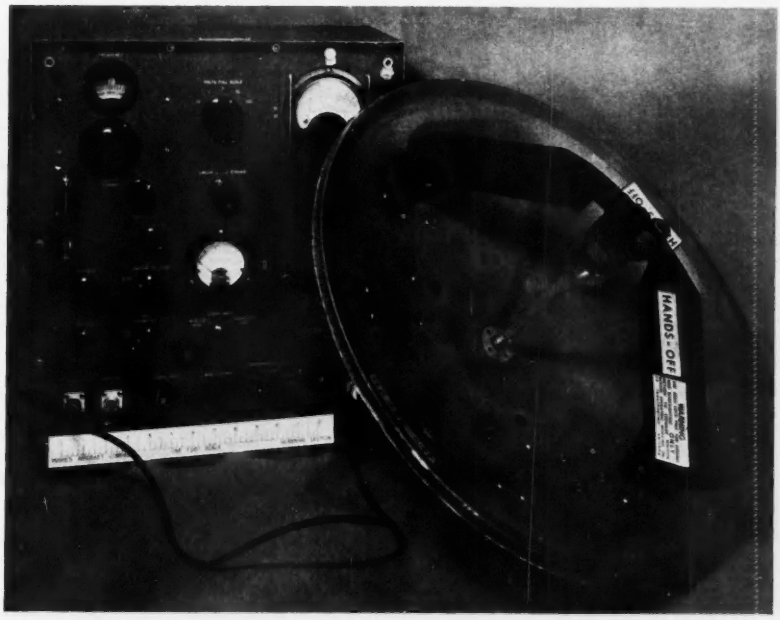


FIG. 12. Conical-lobed antenna and bolometer amplifier.

CONCLUSIONS

The concepts and techniques discussed in this paper represent one of the most interesting and promising new areas of investigation in the field of radiating systems. Without a doubt, the course of electromagnetic systems in future years will be influenced by this philosophy. However, effort in many different directions will be necessary to realize the full capabilities of these methods. Not only will extensive theoretical analysis be necessary on such questions as the accuracy with which the time modulation must be controlled, system analysis, synthesis procedures, etc., but in addition, considerable component research and development is indicated.

The analyses given here have been deliberately chosen for simplicity; however, there is every reason to expect refined techniques to reveal similar results. It is hoped that this discussion will serve to stimulate antenna research in these directions.

REFERENCES

CUTRONA, L. J. 1955. Airborne Battlefield Surveillance Radar Symposium Record, University of Michigan.

SHANKS, H. E. 1958. Hughes Aircraft Company Technical Memorandum No. 585.

TAYLOR, T. T. 1953. Hughes Aircraft Company Technical Memorandum No. 324.

ON THE THEORY OF FLOW OF UNDERGROUND FLUIDS IN COMPRESSIBLE STRATA[1]

A. E. SCHEIDEGGER[2]

ABSTRACT

The problem of flow of a fluid within a compressible porous medium is investigated. It is shown that in general, the motion of the fluid cannot be separated from that of the medium. This leads to a very complex problem of consolidation. However, considerable simplification can be made in applications to the flow of underground fluids. In that case, the general geometry of the consolidation can be predicted since the latter can take place in the vertical direction only. Furthermore, in many cases it is possible to neglect the volume compressibility of the porous matrix.

Two cases have been considered: that of local isotropy of stress and permeability and that of local anisotropy of these two quantities. The basic differential flow equation for the two cases is deduced.

1. INTRODUCTION

Much attention has recently been paid to the flow of fluids through compressible porous media. Fields in which such media are important include not only the study of certain types of petroleum reservoirs but also the investigation of ground-water resources and the problem of filtration through compacting filter cakes.

If one is to treat the problem of flow through compressible porous media fully, one is not able to separate the motion of the fluid from that of the bulk of the porous mass. One has thus a twofold problem: the porous medium is undergoing finite deformations (governed by corresponding equations of motion), and coupled with it, the fluid also is being displaced. All displacements in question are large so that the whole apparatus of finite strain theory has to be invoked. This makes the problem very complicated. Some work on this general problem has been done in connection with studies of consolidation of wet soil (Biot 1941, 1956; Florin 1948); the writer is not aware, however, of an occasion where the full finite strain mechanism would have been introduced and consistently employed. Furthermore, in consolidation studies the flow of fluid contained in the pores is only of minor interest, the flow velocity is usually eliminated altogether, and one is left with equations which refer to the deformation of the porous mass as a whole.

If the flow of the fluid is of primary interest, one must reintroduce it after the problem of consolidation of the medium has been solved, and then solve a new set of differential equations.

In many practical cases, particularly with regard to the study of petroleum reservoirs and other fluid-bearing strata within the earth, it is not necessary

[1]Manuscript received November 27, 1958.
Contribution from the Imperial Oil Research Laboratory, Calgary, Alberta.
[2]Presently Visiting Professor of Geophysics, Seismological Laboratory, California Institute of Technology, Pasadena, Calif., on leave of absence from Imperial Oil Research Laboratory, Calgary.

to treat the problem in its full generality. The displacements due to consolidation are usually small compared with the displacements due to the flow of fluid. Furthermore, it is possible to predict at least the general geometry in which the consolidation will occur so that it is not necessary to solve a complete set of differential equations of motion to determine the latter.

Furthermore, an additional simplification is obtained if it is assumed that the change of the matrix volume as compared with that of the pore volume can be neglected. In many geological applications, this is possible.

Therefore, what is required for the practical study of the flow of fluids in compressible strata is a set of equations of motion for the fluid wherein due allowance is made for the concurrent deformation of the porous medium containing the fluid. There are two possibilities that come to one's mind: either the permeability and the stresses are locally isotropic, or they are not isotropic. One would think, offhand, that it might be possible to treat the porous medium as locally isotropic with regard to permeability and stresses, but recent investigations by Hubbert (1957) show that this is not the case. The possibility of anisotropy occurring must therefore be investigated.

2. STRESSES IN POROUS MEDIA

If porous media are treated as deformable bodies, one is first of all faced with the problem of describing the deformations sustained as functions of the prevailing stresses. The *total stress* in a porous medium (Terzaghi 1951) can be split into two components: one component, termed the *neutral stress*, is the stress in the fluid; the second component, called the *effective stress*, is the difference between the total stress prevailing in the fluid-filled porous medium and the neutral stress. It is this stress which produces the deformation of the porous medium.

In the case that the total stress in the porous medium can be assumed as hydrostatic, this total stress in the porous medium is called *overburden pressure*. The stress state in the fluid-filled porous medium is then completely described by giving any two of the following three quantities: overburden pressure, (π), effective pressure (τ), neutral (fluid) pressure (p). The relation between these three quantities is

$$\pi = \tau(1-P)+pP \tag{2.1}$$

where P is the porosity. The last equation is, in fact, a definition of the "effective" pressure; the latter also can be (and has been) defined differently.

It is clear from the above remarks that, in order to describe the stress state in a fluid-filled porous medium, one needs all in all *seven* parameters: the six parameters of the (symmetric) total stress tensor τ_{ik} and the fluid pressure p. If we have applications to underground strata in mind, it will be possible in most cases to regard the vertical direction as a principal stress direction; the reasons for this are outlined, e.g. in Anderson's (1942) book. The other two principal directions, then, lie somewhere in the horizontal plane.

Using a Cartesian co-ordinate system x, y, z with z denoting the vertical and x, y principal axes, one has, under the present simplifying assumptions,

three significant stress components: τ_{xx}, τ_{yy}, and τ_{zz}; the other stress components, viz. τ_{xy}, τ_{xz}, and τ_{yz} are zero. The seven parameters, therefore, reduce to four (viz. τ_{xx}, τ_{yy}, τ_{zz}, and p). A further simplification is achieved if one assumes the stress state as isotropic; one then has only *two* significant parameters (overburden pressure and fluid pressure).

The next task is to describe the deformation of the porous medium during compaction by some proper means. To this end, we follow the general method in finite strain theory (cf. Scheidegger 1956) of introducing "parameters" ξ_α. These parameters are the Cartesian co-ordinates of the porous medium in an arbitrary state (e.g. under zero overburden stress and zero fluid pressure). The parameters ξ_α will be the variables to which everything will ultimately be referred.

Let the co-ordinates of the point ξ_α in the stressed state be $x_i(\xi_\alpha)$. Then we have

$$x_i = x_i(\xi_\alpha, \tau_{ik}, p, P), \tag{2.2}$$

$$\xi_\alpha = \xi_\alpha(x_i, \tau_{ik}, p, P) \tag{2.3}$$

where it is implied that the rheological condition of the porous medium is such that the latter's deformation is entirely given if a certain value is assumed for the total stress τ_{ik}, the fluid pressure p, and the fluid content (volume porosity P). This, of course, neglects all the inertia forces connected with the motion of the medium. However, we allow for the possibility of hysteresis by specifying single valuedness for the x's only with regard to a certain direction from which the stress state in question was attained.

With the above equipment, one is able to proceed to the hydrodynamics of the fluid contained in the pores.

3. THE HYDRODYNAMIC EQUATIONS

We now proceed to set up the hydrodynamic equations of the fluids. The fluid will be connected with the porous medium during the latter's consolidation. What is of ultimate interest is only the motion of the fluid in terms of the parameters ξ_α. As is well known (cf. the writer's book (Scheidegger 1957)), one needs three sets of conditions to make any flow problem determined: (i) Darcy's law, (ii) the constitutive equation of the fluid, and (iii) the continuity condition. The flow equations should be set up in such a fashion that the various quantities occurring therein can be conveniently measured.

Beginning with *Darcy's law*, we *assume* that it can be written as follows

$$q_\alpha = -\frac{k_{\alpha\beta}}{\mu}\frac{\partial p}{\partial \xi_\beta} \tag{3.1}$$

where the summation convention is used. The quantity q_α is the filter velocity vector as referred to the filtration surface in the zero strain state, μ the viscosity, and p the pressure of the liquid. The quantity $k_{\alpha\beta}$ is the (*variable*) permeability tensor; there is some ambiguity as to whether or not it would have to be taken inside the gradient. This question, however, also occurs in

the discussion of relative permeability and a dissertation of why the form given above should be adopted over other possibilities may be found on p. 62 of the writer's book (Scheidegger 1957).

The *constitutive equation* of the fluid is simply

$$\rho = \rho(p) \tag{3.2}$$

where ρ is the density, p the pressure.

This is the same constitutive equation as that encountered in flow through incompressible porous media.

Finally, we turn our attention to the *continuity condition*. Expressing the condition that the mass flow out of a small volume of $d^3\xi$ of porous medium (with vector-surface elements dA_α) must be equal to the change of mass inside yields

$$\oint \rho q_\alpha \, dA_\alpha = -\frac{\partial}{\partial t}\int_v (1+\kappa)\rho P d^3\xi \tag{3.3}$$

where P is the porosity at the point ξ, and $(1+\kappa)$ is the volume factor induced by the consolidation of the porous medium. In virtue of Gauss's theorem, the last equation can be written as follows:

$$\int \frac{\partial}{\partial \xi_\alpha} \rho q_\alpha \, d^3\xi = -\int \frac{\partial}{\partial t}\left[(1+\kappa)\rho P\right] d^3\xi \tag{3.4}$$

or

$$-\frac{\partial}{\partial \xi_\alpha}(\rho q_\alpha) = \frac{\partial}{\partial t}\left[(1+\kappa)\rho P\right]. \tag{3.5}$$

This is the continuity condition.

Combining the three basic equations, one ends up with an equation for p only.

$$\frac{\partial}{\partial t}\left[(1+\kappa)\rho P\right] = \frac{\partial}{\partial \xi_\alpha}\left[\frac{\rho}{\mu} k_{\alpha\beta} \frac{\partial p}{\partial \xi_\beta}\right]. \tag{3.6}$$

This is the fundamental hydrodynamic equation.

Equation (3.6) can be further simplified if it can be assumed that the total change of bulk volume in any one specimen is entirely produced by change of the pore volume, i.e. if it can be assumed that the volume of the solid matrix changes only negligibly. Then, one can write

$$\text{bulk volume } (p, \pi) = \text{bulk volume}_0\,(1+\kappa(p, \pi)), \tag{3.7}$$

$$\text{pore volume } |_0 = P_0 \text{ bulk volume}_0, \tag{3.8}$$

$$P = \frac{\text{pore volume } (\pi, p)}{\text{bulk volume } (\pi, p)} = \frac{\text{pore volume } (\pi, p)}{\text{bulk volume}_0\,(1+\kappa(p, \pi))}, \tag{3.9}$$

but:

$$\left\{\begin{aligned} &\text{pore volume } (\pi, p) - \text{pore volume } (0) \\ &= \text{bulk volume } (\pi, p) - \text{bulk volume } (0) \\ &= \text{bulk volume}_0\,[(1+\kappa(p, \pi)) - 1] = \kappa(p, \pi) \text{ bulk volume}_0; \end{aligned}\right. \tag{3.10}$$

furthermore:

(3.11) pore volume (π, p) = bulk volume$_0$ $(\kappa(p, \pi))$+pore volume$_0$,

(3.12) pore volume (π, p) = bulk volume$_0$ $(P_0+\kappa(p, \pi))$;

hence:

(3.13) P = [bulk volume$_0$ $(P_0+\kappa(p, \pi))$]/[bulk volume$_0$ $(1+\kappa(p, \pi))$],

or

(3.14) $P = [P_0+\kappa(p, \pi)]/[1+\kappa(p, \pi)]$.

Therefore, in the simplified case at present under consideration, the fundamental hydrodynamic equation reduces to

$$\frac{\partial}{\partial t}[\rho(P_0+\kappa(p, \pi))] = \frac{\partial}{\partial \xi_\alpha}\left(\frac{\rho}{\mu} k_{\alpha\beta} \frac{\partial p}{\partial \xi_\beta}\right). \tag{3.15}$$

4. THE CASE OF LOCALLY ISOTROPIC STRESS AND PERMEABILITY

Let us consider first the case of locally isotropic stress and permeability. In that case, according to earlier remarks, the stress state is described by two parameters, viz. by the fluid pressure p and by the overburden pressure π. The various parameters required in the flow equation (3.6) can be determined as follows:

(*a*) *Permeability*

Let a unit (at zero strain) volume be enclosed in a *pressure cell*. The parameters ξ_α then correspond to this state. Then, the pressure is raised to π, p and one measures a certain flow. Let the volume reduction be by the relative amount $-\kappa(\pi, p)$, i.e. the length reduction by $-\frac{1}{3}(\kappa, p)$ (neglecting squares of κ). What is measured is a superficial permeability k_{sup}, according to the formula

$$\frac{Q}{A_0} = \frac{k_{\text{sup}}}{\mu}\frac{p_2-p_1}{L_0} \tag{4.1}$$

where the index "0" stands to indicate that the original (zero strain) area and length are used to calculate k_{sup}. The true permeability of the sample, assuming cubic volume change, is given by

$$\frac{Q}{A_0(1+2\kappa/3)} = \frac{k_{\text{true}}}{\mu}\frac{p_2-p_1}{L_0(1+\kappa/3)} \tag{4.2}$$

hence

$$k_{\text{true}} = k_{\text{sup}}\frac{1+\kappa/3}{1+2\kappa/3} = \frac{3+\kappa}{3+2\kappa}k_{\text{sup}} = \left(1-\frac{\kappa}{3}\right)k_{\text{sup}} \tag{4.3}$$

where the last equal sign has been obtained upon neglecting high order terms.

However, the volume change in the *field* is not cubic, although it is cubic in the pressure cell. According to earlier remarks, the compaction occurs in the vertical (ζ) direction only. Assuming the true permeability under pressure to be the same in the field as in the pressure cell, we can express the con-

dition for the "superficial" flow q_{sup}, i.e. the flow in terms of zero strain co-ordinates of the field. We have

$$q_{\text{sup}\,x} = \frac{Q_{x(\text{true})}}{L_0^2} = \frac{1}{L_0^2}\frac{p_2-p_1}{L_0}\frac{k_{\text{true}}}{\mu}L_0^2\,(1+\kappa) \tag{4.4}$$

and hence

$$k_{\xi\xi} = k_{\eta\eta} = (1+\kappa)\,k_{\text{true}} = (1+\kappa)\,\frac{1+\kappa/3}{1+2\kappa/3}\,k_{\text{sup}}. \tag{4.5}$$

Neglecting squares of κ:

$$k_{\xi\xi} = \frac{3+4\kappa}{3+2\kappa}\,k_{\text{sup}} = \left(1+\frac{2\kappa}{3}\right)k_{\text{sup}}. \tag{4.6}$$

Similarly

$$q_{\text{sup}\,z} = \frac{Q_{z(\text{true})}}{L_0^2} = \frac{1}{L_0^2}\frac{p_2-p_1}{L_0(1+\kappa)}\frac{k_{\text{true}}}{\mu}L_0^2 \tag{4.7}$$

and hence

$$k_{\zeta\zeta} = \frac{1}{1+\kappa}\,k_{\text{true}} = \frac{1}{1+\kappa}\,\frac{1+\kappa/3}{1+2\kappa/3}\,k_{\text{sup}}. \tag{4.8}$$

Neglecting squares of κ:

$$k_{\zeta\zeta} = \frac{1+\kappa/3}{1+5\kappa/3}\,k_{\text{sup}} = \frac{3+\kappa}{3+5\kappa}\,k_{\text{sup}} = \left(1-\frac{4\kappa}{3}\right)k_{\text{sup}}. \tag{4.9}$$

The permeability tensor in the system ξ, η, ζ (ζ vertical) has therefore the following form

$$k_{\alpha\beta} = \begin{pmatrix} 1+\frac{2}{3}\kappa & 0 & 0 \\ 0 & 1+\frac{2}{3}\kappa & 0 \\ 0 & 0 & 1-\frac{4}{3}\kappa \end{pmatrix}. \tag{4.10}$$

It is therefore seen that, in order to determine $k_{\alpha\beta}$ one needs *one* permeability measurement (k_{sup}) and *one* measurement of the bulk compressibility for each possible pair of variables p, π. During the depletion of underground strata of their fluid content, it can be assumed that the overburden pressure stays constant; the measurements of k_{sup} and κ have therefore to be made with given overburden pressure π for all values of p.

(*b*) *Porosity*

Another parameter that enters into eq. (3.6) is the combination $(1+\kappa)\rho P$. It is customary to measure the differential quotient $(1/P_0)(d/dp)[(1+\kappa)P]$ where the index 0 refers to the zero strain state. This differential quotient

has been called "effective compressibility" (McLatchie, Hemstock, and Young 1958). If the latter be measured for one series of values π, p arriving at the field overburden pressure and fluid pressure from the zero strain state, and then for the given overburden pressure for all pertinent values of the fluid pressure p, and furthermore, if the rheological equation of the fluid $\rho(p)$ be known, then the required values of $(1+\kappa)\rho P$ can be obtained by integration. It is then possible to eliminate all unknowns except p from (3.6) and hence, to solve any underground flow problem.

5. ANISOTROPIC STRESS AND PERMEABILITY

It is interesting to note the modifications that are required in the above scheme if one is to generalize it to take care of anisotropic stresses and permeability. The general equation of flow (3.6) is, of course, still applicable, but different means have to be found to measure the required quantities.

In order to make the problem amenable to investigation at all, it is necessary to assume that the vertical direction is a principal direction for stress and for permeability. The other principal axes then lie somewhere in the horizontal plane.

In order to measure the quantities required in equation (3.6), a means must be found to restore in the laboratory the state in which the rock was in the field. In general, this requires a separation of the three principal stresses from one another. Thus, a piece of rock being taken from a well will, first of all, expand *cubically* (because it is taken out of confinement) while it is in transit to the laboratory. Then it must be enclosed in some device in which the total stresses τ_{zz} and τ_{xx}, τ_{yy}, as well as the fluid pressure p can be brought up to the values which prevailed in the field. This reproduces the original state of the rock. For the following, it is useful to refer the parameters ξ_α to this state rather than to the zero strain state.

Afterward, during a test run, τ_{zz} must be held constant, the lateral stresses (τ_{xx}, τ_{yy}) must be adjusted in such a manner that there is no lateral motion of the rock (which is the boundary condition prevailing in the field) and the vertical (k_z) and horizontal (k_x, k_y) permeabilities must be measured separately for all values of the fluid pressure p. This yields immediately the permeability tensor $k_{\alpha\beta}$, viz.

$$k_{\alpha\beta} = \begin{pmatrix} k_x & 0 & 0 \\ 0 & k_y & 0 \\ 0 & 0 & k_z \end{pmatrix} \tag{5.1}$$

if ξ_α, the lengths, and the areas are referred to the original field state as defined above (rather than the zero strain state). The other quantity that is required in the flow equation is $[(1+\kappa)\rho P]$; it can be measured just as in the isotropic case. Since the ξ_α now refer to the field state, the effective compressibility must be measured with this field state as basic state (P_0 referring to this state). It may be noted that it is now no longer necessary to obtain the bulk compressibility and the "effective" (pore) compressibility separately, because an additional permeability measurement has been substituted for same.

Again, with the various quantities measured, all unknowns except p can be eliminated in the fundamental flow equation (3.6) and hence it is possible to solve any underground flow problem.

6. DISCUSSION

The preceding analysis enables one to draw some interesting conclusions with regard to practical requirements in petroleum engineering and in hydrological studies. These are:

(i) The problem of flow of fluids through consolidating porous media is quite complicated in its full generality, but for practical application to the study of the flow of underground fluids, substantial simplifications are permissible. One thus can assume that the compaction takes place in the vertical direction only.

(ii) If it is further assumed that the stress state in the porous medium is hydrostatic, and the permeability locally isotropic, then the basic differential equation for the flow contains *three* independent quantities (in addition to the pressure–density relationships of the fluid) that have to be measured as functions of the pressure and overburden pressure. These are (for instance) the bulk-volume reduction κ of the medium, the superficial permeability k_{sup}, and the effective compressibility $(1/P_0)d[(1+\kappa)P]/dp$, the latter being defined as the change of pore volume per unit (zero strain) pore volume contained in the medium per unit (fluid) pressure change. Furthermore, the pressure–density relationship for the fluid must be known. It may be noted that the effective compressibility is not required for the solution of steady-state problems.

(iii) If it is conceded that the permeability as well as the stress may be different in the horizontal direction from that in the vertical direction, then one can note the following facts:

(*a*) The variables are now *four* instead of two: the fluid pressure p (just as before) and *three* values of the overburden *stress* $\tau_{\xi\xi}$, $\tau_{\eta\eta}$, and $\tau_{\zeta\zeta}$ (instead of the overburden *pressure* π).

(*b*) There are now *seven* quantities instead of three quantities that have to be measured: (i) the effective compressibility just as before; (ii) *three* permeability values (in the horizontal and two principal vertical directions) instead of one as before, and (iii) *three* quantities (change of length in the vertical and in the two principal horizontal directions, respectively) which are taking the place of the bulk compressibility. If it is again assumed that compaction in the field can occur in the vertical direction only, then the last three quantities reduce to one (to the compaction in the vertical direction), which can be omitted entirely if flow problems only are of interest as distinct from the compaction of the medium, but the stress states that are permissible are then only those which do not produce any lateral motion.

With the required quantities known as functions of pressure and overburden stress, it is in principle possible to solve any underground flow problem.

It remains to compare the above investigations with other recent attempts

at elucidating the effect of fluid pressure decline on the volume changes of porous rocks. It may be noted that Geertsma (1957) noted that two different compressibilities have to be measured to describe properly the volume changes of a porous medium under compression. However, Geertsma does not proceed to write down the equations of motion. He therefore introduces "viscous" deformations of the medium. It is the contention of the writer that these "viscous" deformations are the automatic outcome of the general theory of consolidation.

Similarly, Barenblatt and Krylov (1955) wrote down an equation of flow in fluids in deformable porous media. In essence, their equation is expressed in terms of partial derivatives of the porosity with regard to the fluid pressure (the mean effective pressure being held constant) and with regard to the mean effective pressure (the fluid pressure being held constant). This, of course, yields an equivalent expression to that given in this paper, except that the choice of variables makes it rather difficult to adapt the formulas to practical cases.

ACKNOWLEDGMENTS

The writer wishes to acknowledge that he profited much from discussions with Charles Mack of Imperial Oil Limited, in Sarnia. Dr. Mack pointed out the importance of anisotropic stress states, for which the writer wishes to thank him. The work reported in the present paper was done at the Research Laboratory of Imperial Oil Limited in Calgary, under the leadership of James W. Young. Without the stimulating atmosphere at the laboratory, this would not have been possible. The writer also wishes to express his gratitude to Imperial Oil Limited for the permission granted to publish this study.

REFERENCES

ANDERSON, E. M. 1942. The dynamics of faulting and dyke formation with applications to Britain (Oliver and Boyd, Ltd., Edinburgh).
BARENBLATT, G. I. and KRYLOV, A. P. 1955. Izvest. Akad. Nauk S.S.S.R. Otdel. Tekh. Nauk **2**, 5.
BIOT, M. A. 1941. J. Appl. Phys. **12**, 426.
——— 1956. J. Appl. Phys. **27**, 459.
FLORIN, V. A. 1948. Doklady Akad. Nauk S.S.S.R. **57**, 219.
GEERTSMA, J. 1957. Trans. AIME **210**, 331.
HUBBERT, M. K. 1957. Trans. AIME **210**, 153.
MCLATCHIE, A. S., HEMSTOCK, R. A., and YOUNG, J. W. 1958. J. Petrol. Technol. **10** (6), 49.
SCHEIDEGGER, A. E. 1956. Can. J. Phys. **34**, 498.
——— 1957. The physics of flow through porous media (University of Toronto Press).
TERZAGHI, K. 1951. Theoretical soil mechanics (Chapman and Hall, Ltd., London).

AN EXPERIMENTAL STUDY OF INTERCEPTION NOISE IN ELECTRON STREAMS AT MICROWAVE FREQUENCIES[1]

B. A. McIntosh[2]

ABSTRACT

The noise added to an electron stream by the interception of a fraction of the current has been studied experimentally at a frequency of 3 kMc/s. The electron beam was produced in a demountable vacuum system by a parallel-flow Pierce gun in a confining magnetic field. A series of circular apertures and mesh grids on a plate capable of being moved within the vacuum chamber allowed the interception of various fractions of the total beam current. The excess noise caused by interception was measured at the anode of the electron gun and at various points in a drift region.

Interception noise caused by mesh grids was found to be of much greater magnitude than that caused by circular apertures. The absolute level of the excess noise and its variation with the strength of the confining magnetic field were found to be in reasonable agreement with existing theory as modified by the author.

It was shown that current interception excites a standing wave of noise along the electron beam in the drift region. Measurements made on a temperature-limited electron beam have shown that interception noise can arise in a region where noise smoothing is a consequence of processes external to the electron gun.

I. INTRODUCTION

It has long been known that if a portion of the electron current in a vacuum tube is intercepted by a structural element, the inherent noise in the remaining portion of the electron stream is increased. Theoretical predictions of the increased noise in multigrid tubes as opposed to triodes were verified experimentally (North *et al.* 1940). These measurements were made on common radio receiving tubes at low frequencies.

As a result of the recent development of beam-type amplifier tubes for use at microwave frequencies, the role of interception noise in determining the noise properties of such tubes is of interest. Excess noise caused by current interception has been observed experimentally in microwave tubes (Cutler and Quate 1950; Fried and Smullin 1954). These workers noted that the interception of a cylindrical beam by a circular aperture caused much less excess noise than would be predicted by North's theory, but comprehensive experimental measurements have been lacking. North's theory was thought to be inapplicable to microwave tubes since it was based on a particular mechanism of noise smoothing which was believed to be invalid at high frequencies. North also assumed that current was intercepted by a fine mesh grid.

A significant contribution to the theory of interception noise has been made by Beam (1955). He calculated the excess noise caused by interception

[1]Manuscript received November 10, 1958.
Contribution from the Eaton Electronics Research Laboratory, McGill University, Montreal, Que.
[2]Now with the Radio and Electrical Engineering Division, National Research Council, Ottawa, Canada.

of an electron beam by a circular aperture for the case of a beam collimated by an axial magnetic field. Since most microwave tubes require a magnetic field to confine the long electron beams used, such theory has considerable practical importance. Beam presented experimental measurements to verify the functional variation of interception noise with magnetic field strength but made no attempt to verify predictions concerning the magnitude of the excess noise.

It is well known that when an electron beam drifts at constant velocity in a field-free region a standing wave of noise exists on the beam. Robinson (1954) postulated that interception would excite a noise standing wave but experimental verification of this phenomenon has not been reported in the literature. A knowledge of the behavior of the interception-noise wave in the drift space is of importance in minimizing the noise figure of beam-type tubes such as the travelling-wave tube.

It has been pointed out by Beam (1955) that a theory of interception noise based on random probability of interception of electrons leads to the same result as North's theory yet requires no assumptions concerning the mechanism of smoothing as did North's theory. Experimental verification was again lacking.

II. THEORY

Although this research has been primarily of an experimental nature, some attempt has been made to clarify and extend existing theories. North's theory of interception noise (North *et al.* 1940) is based on his conception of the smoothing action of the potential minimum which lies immediately in front of the electron-gun cathode when the emission is space-charge limited. Fluctuations in the total emission current were believed to be reduced in magnitude by "compensating pulses" of current produced by the potential minimum. Interception of a fraction of the current by a mesh grid destroys the correlation between the initial fluctuations and the compensating fluctuations and thereby caused the noise to be increased. The validity of this theory was therefore in doubt when it was found that the potential minimum played little part in the noise-smoothing process at microwave frequencies.

Robinson and Kompfner (1951) have given a simplified derivation of North's result, which may be written in the form

$$\overline{i_{2n}^2} = \Gamma_1^2 k_0 2eI_{01}\Delta f + (1-\Gamma_1^2)(1-k_0)k_0 2eI_{01}\Delta f \tag{1}$$

where $\overline{i_{2n}^2}$ is the mean square of the noise-current fluctuations that exist in the beam after it has been intercepted;

e is the electron charge;

I_{01} is the direct current before interception;

k_0 is the fraction of the current I_{01} which is transmitted through the intercepting grid;

Δf is the frequency bandwidth in which the noise is observed;

Γ_1^2 is called the "smoothing factor" of the electron beam before interception.

The smoothing factor may be defined as the ratio of the actual noise power in the beam to full shot-noise power in the same direct current; that is,

$$\Gamma_1^2 = \frac{\overline{i_{1n}^2}}{2eI_{01}\Delta f}.$$

(The subscript $_1$ is used to denote quantities before interception; $_2$ for quantities after interception.) The smoothing factor of the beam after interception may be written

$$\Gamma_2^2 = \frac{\overline{i_{2n}^2}}{2eI_{02}\Delta f} = \Gamma_1^2 + (1-\Gamma_1^2)(1-k_0). \tag{2}$$

It has been demonstrated by Beam (1955) that the theory of interception noise can be based solely on the assumption of random interception of electrons without recourse to any assumptions concerning the nature of the smoothing process. North's result is in fact applicable to high-frequency tubes provided the intercepting electrode is a fine mesh grid, which is to say that the probability of interception of electrons must be the same at all points on a cross section of the electron stream. The author has shown (McIntosh 1958) that a criterion for the necessary "fineness" of mesh may be set up in terms of the concept of "mean spreading radius" of electrons. This is essentially the radial distance about any point on a cross section of the electron stream, over which the probability of random interception is uniform; it is determined by the distribution of random transverse velocities of the electrons. If the dimensions of the grid openings are larger than the mean spreading radius, as in the case of a single circular aperture, the effective area over which interception of electrons will be random is much smaller than the total beam area. The excess noise added to the beam is therefore much smaller than the North value.

The probability that an electron at radius r in the beam be transmitted through an aperture of radius r_a has the form sketched in Fig. 1(a). The region in which transmission or interception may be random is essentially determined by the product $k_0(r)[1-k_0(r)]$ shown in Fig. 1(b). If these factors are averaged over the total beam area they yield respectively the fraction of current transmitted and the fraction of the current in which interception is random. Following Beam's notation (Beam 1955), these are

$$\Omega_1 = \frac{\int J(r)k_0(r)dA}{\int J(r)dA}, \tag{3}$$

$$\Omega_2 = \frac{\int J(r)k_0(r)[1-k_0(r)]dA}{\int J(r)dA}. \tag{4}$$

$J(r)$ is the direct current density which in general may be a function of transverse co-ordinates. The integration is over the entire beam area.

The smoothing factor after interception may then be written

$$\Gamma_2^2 = \Gamma_1^2 + (1-\Gamma_1^2)\frac{\Omega_2}{\Omega_1}. \tag{5}$$

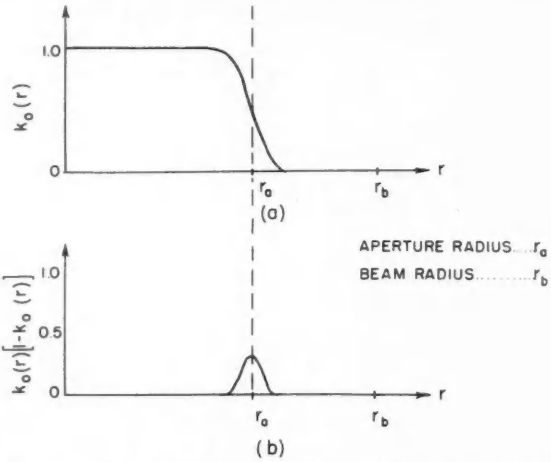


FIG. 1. (*a*) The probability that an electron associated with a point at distance r from the center of the beam will be transmitted through an aperture of radius r_a.
(*b*) A plot of the region, near the aperture edge, in which interception or transmission of electrons is random.

Evaluation of Ω_1 *and* Ω_2

To simplify calculation, it may be assumed that the current density $J(r)$ is constant out to the beam edge at $r = r_b$ and is zero thereafter. Then for all practical purposes,

$$\Omega_1 = \left(\frac{r_a}{r_b}\right)^2 \quad , \quad r_a \leqslant r_b \, . \tag{6}$$

In cases where r_a is nearly equal to r_b, Ω_1 is approximately unity but $1-\Omega_1$ is of significance and may be calculated from the equations given by Beam.

The author has shown (McIntosh 1958) that an approximate expression for Ω_2 is

$$\Omega_2 = \pi r_a/8ar_b^2\{1+\mathrm{erf}[2(\pi)^{-\frac{1}{2}}a(r_b-r_a)]\} \tag{7}$$

where "erf" stands for "error function of". The parameter a is the inverse of the mean spreading radius.

When the electron beam is confined by a magnetic field, transverse thermal velocities cause the electrons to execute spiral trajectories. The Maxwellian distribution of spiral radii determines the mean spreading radius as

$$\frac{1}{a} = \frac{(2mkT_c)^{\frac{1}{2}}}{Be} \tag{8}$$

where m is the mass of the electron,
k Boltzmann's constant,
T_c the cathode temperature,
B is the strength of the confining magnetic field.

If no magnetic field exists, the mean spreading radius is determined by the transit time τ from the cathode to the plane of interception. For a parallel-flow, Pierce-type electron gun there results

$$\frac{1}{a} = \tau\left(\frac{2kT_c}{m}\right)^{\frac{1}{2}}. \tag{9}$$

Of most interest to the designer of beam-type tubes is the magnitude of interception noise for small amounts of current intercepted by an aperture near the beam edge. In this case it is unrealistic to assume that the beam has a rectangular current-density profile since, with a finite confining field, the spiralling of electrons near the edge of the beam tends to "smear out" the edge. It may be shown (McIntosh 1958) that the radial dependence of current density is given by

$$J(r) = \tfrac{1}{2}J_0\{1-\mathrm{erf}[a(r_b-r)]\} \tag{10}$$

where J_0 is the average current density given by

$$J_0 = \frac{I_{01}}{\pi r_b^2}.$$

Using this expression, modified values of Ω_2 have been calculated and are compared in Table I with values given by Beam (1955).

TABLE I

Calculated values of Ω_2 modified to take into account radial variation of current density

ar_a	$(r_a/r_b)-1$	Ω_2 modified	Ω_2 Beam
50	-0.01	1.0×10^{-2}	1.3×10^{-2}
	0	0.59	0.80
	0.01	0.24	0.35
20	-0.025	2.5	2.9
	0	1.5	2.0
	0.025	0.59	0.80
10	-0.05	5.0	6.0
	0	2.9	3.9
	0.05	1.2	1.7

Although this first-order correction yields a reduction in Ω_2 of only 1 to 2 db, it does indicate that the thinning-out of current density near the beam edge tends to reduce the effect of interception.

Application to Mesh Grids

Beam's theory may be extended to mesh grids in the following manner. The electron beam may be divided into a bundle of smaller beams each of which is associated with one opening in the grid. The fraction of current intercepted from one of the small beams is equal to the fraction intercepted from the whole stream provided the current density is reasonably uniform over each of the small beams. A smoothing factor Γ_2^2 may be calculated for one of the small beams in the same manner as for a single aperture, and

since Γ_2^2 is a measure of noise per unit current, the value thus calculated applies to the whole beam. This analysis neglects possible interaction among the small beams, i.e., the fact that the small beams have no "edge", and that the radii of spiralling electrons may be large enough to overlap several grid modules. Such effects should be small if the mean spreading radius $1/a$ is less than the radius of one of the grid openings. When $1/a$ becomes greater than this radius, the excess noise will approach the North value.

III. APPARATUS

The Mechanical System

Figure 2 (*a*) is a photograph of the apparatus used to make interception-noise measurements. The important and central portion of the apparatus is a vacuum chamber, maintained at a pressure of 3×10^{-6} mm of mercury, in which an electron gun and an intercepting electrode could be moved longitudinally with respect to a resonant cavity mounted on the right end of the long brass chamber. The large solenoid, mounted on wheels and capable of being rolled over the vacuum chamber, served to collimate the electron beam. The rack-like table was used to mount electronic apparatus.

Figure 2(*b*) is an exploded view of the components within the vacuum chamber; the outer vacuum jacket has been removed. The electron gun and intercepting electrodes were mounted on frames which slid on teflon inserts along the inner surface of the brass tube seen at right. Over a longitudinal distance of 40 centimeters in which the motion took place, the lateral deviation of the sliding structures was less than 0.07 millimeters.

The intercepting electrodes used were three apertures—of diameters .080 in., .060 in., .040 in.—and two mesh grids. The grids were a standard type used by Varian Associates in the manufacture of klystrons. One of the grids was mounted perpendicular to the beam axis, the other at an angle of 20° in order to intercept a larger fraction of the beam current. All electrodes were assembled on a single plate which pivoted on the plunger controlling the longitudinal motion. By rotating the plunger, any one of the electrodes could be moved into the path of the beam.

The resonant cavity was a standard re-entrant type with a resonant frequency of 3050 Mc/s. Mounted directly behind the cavity, on the outside of the circular brass plate seen in Fig. 2(*b*), is the electron collector. It was surrounded by a soft-iron magnetic shield in order to eliminate the spurious noise caused by reflected secondary electrons passing back through the cavity gap (Kornelsen 1957).

The electron gun initially used in the noise measurement followed a design originated by Kornelsen (1957) especially for use in demountable vacuum systems. It employed an indirectly heated tantalum emitter which was not subject to contamination at the moderate vacuum (10^{-5} to 10^{-6} mm Hg) attainable in a demountable system. Nor was it affected by repeated exposure to atmospheric pressure when changes had to be made in the system. For measurements of the variation of interception noise with magnetic field strength, an electron gun with an oxide-coated cathode was used since the

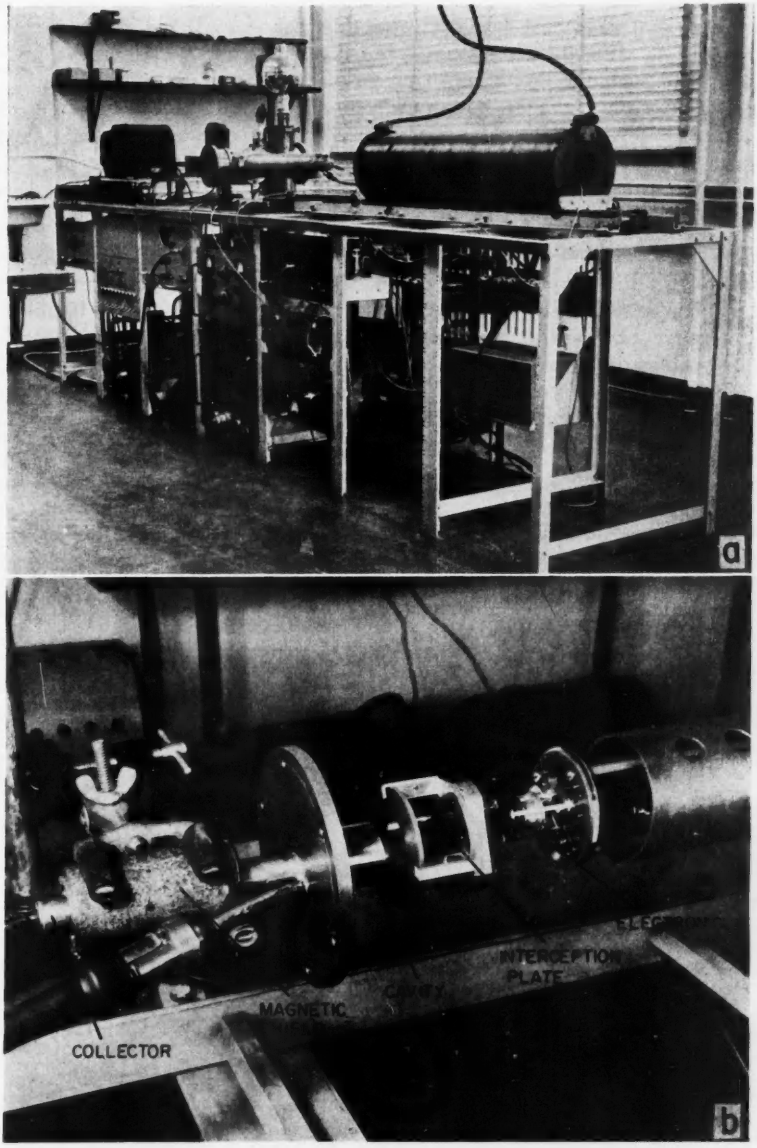


FIG. 2. (*a*) A photograph of the apparatus used to measure interception noise.
(*b*) An exploded view of the internal components of the tube structure.

bombarded-cathode gun did not operate satisfactorily in magnetic fields of less than 700 gauss. At low field strengths the bombarding beam was so divergent that insufficient current was delivered to the tantalum button to achieve space-charge limitation in the primary gun.

Except for the cathodes, both guns were of identical, parallel-flow, Pierce-type design (Pierce 1949). The experimental perveance was 1.1×10^{-7} (amp) $(\text{volt})^{-3/2}$. Cathode temperatures for space-charge-limited emission were 2300° K for the tantalum cathode and 1350° K for the oxide cathode.

The Measuring System

The basic components of the measuring system are shown in Fig. 3. The electron beam was gated by a symmetrical square wave at 35 c/s. Noise power coupled from the electron beam to the resonant cavity was amplified by a conventional superheterodyne receiver and, following audio detection and amplification, was detected coherently. McFarlane (1958) has analyzed this system and shown that the limiting sensitivity is a signal 35 db below the self-noise of the receiver. The measurement of shot-noise power in a beam current of 1 μa has been readily achieved.

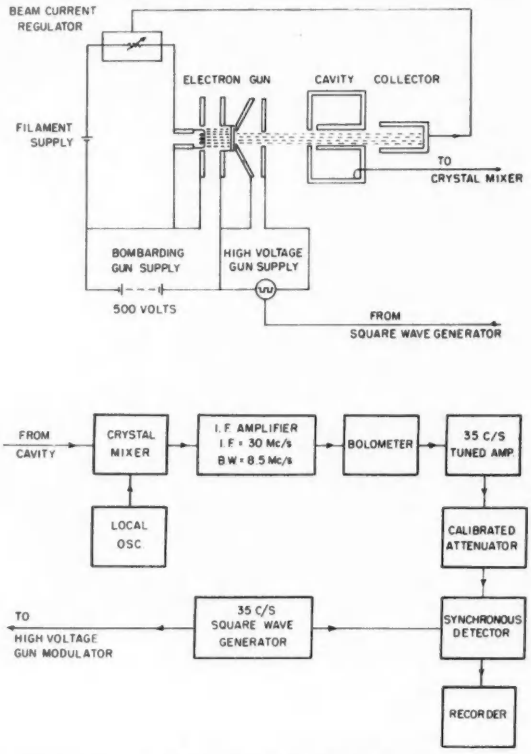


FIG. 3. The electronic measuring apparatus.

IV. MEASUREMENTS

In the first series of measurements for which the apparatus was used, the interception plate and the cavity were located at the anode of the electron gun. Noise smoothing, before and after interception by the various electrodes, was obtained as a function of electron gun voltage at a fixed value of magnetic field strength. The variation of smoothing with magnetic field strength was measured for a constant value of gun voltage.

The investigation of the effect of interception on the noise space-charge waves in the drift space consisted in measurements of noise power as the electron gun was moved back from the cavity, with the interception plate at various positions between gun and cavity. These measurements were carried out for the condition of space-charge-limited emission and for temperature-limited emission.

Calibration of the Measuring System

The apparatus was calibrated by a method used by Cutler and Quate (1950). With the cavity at the anode of the electron gun, relative noise power was measured as the beam current was increased from zero to its maximum value at space-charge limitation. Typical curves of noise power versus collector current which resulted from measurements using the bombarded-cathode gun are shown in Fig. 4. In the region of temperature-limited current the noise power rose linearly with collector current as predicted theoretically by the shot-noise relation,

$$\overline{i_n^2} = 2eI_0\Delta f.$$

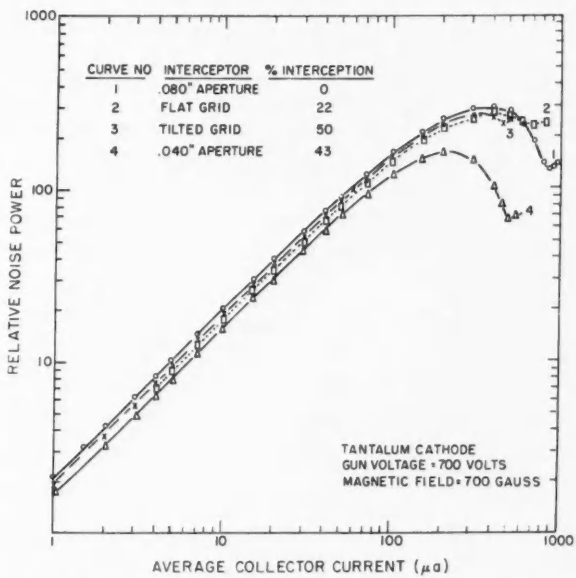


FIG. 4. Typical curves of the variation of noise power with collector current.

The onset of smoothing, as the cathode temperature was increased further, may be clearly seen from the curves. At the maximum value of space-charge-limited current, the ratio of measured noise power to the equivalent shot-noise power determined by extrapolating the linear portion of the curve to the same current yields the smoothing factor Γ^2.

Measurements of the variation of interception noise with magnetic field strength were normalized to shot noise by means of smoothing curves obtained for a fixed value of magnetic field. Similarly, in the case of measurement of noise space-charge waves along the electron beam, the relative noise power levels of the maxima and minima were referred to the level at the anode which had been determined with respect to shot noise from the smoothing measurements.

V. EXPERIMENTAL RESULTS AND DISCUSSION

(*a*) *Interception Noise as a Function of Magnetic Field Strength*

The measurements to be described here were made with the emission from the oxide-coated cathode space-charge-limited at an anode voltage of 700 volts.

Figure 5 shows the measured values of beam noise after interception as a

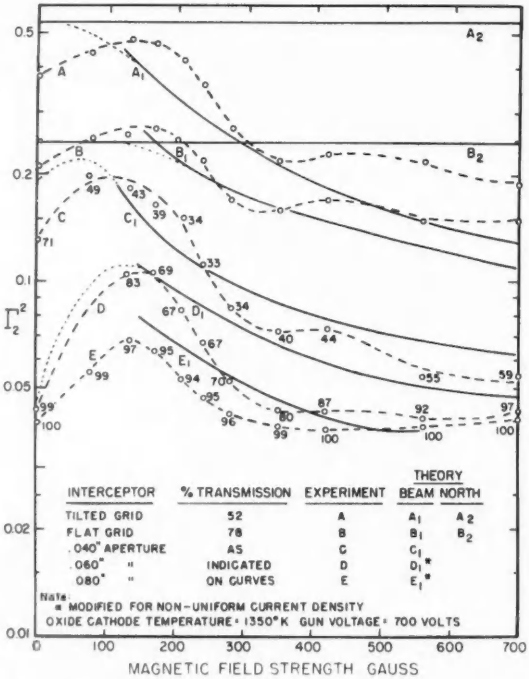


FIG. 5. The variation of noise smoothing after interception as a function of the strength of the confining magnetic field. Ordinate values are noise power normalized to full shot noise in the transmitted current.

function of the strength of the confining magnetic field. Ordinate values represent the noise content of the transmitted beam normalized to full shot noise in the transmitted current. In the case of the circular apertures, the amount of current transmitted was a strong function of the value of the magnetic field. The number associated with each experimental point in Fig. 5 is the per cent transmission at that field strength. As noted on the graph, the transmission factor of the grids was independent of magnetic field strength.

The striking difference between the amount of noise added to the beam by grid interception and that added by aperture interception is evident from these measurements. For comparable fractions of current intercepted the magnitude of the smoothing is from 3 to 6 db less for grid interception than for aperture interception.

For the purpose of comparison with theory, the values of the smoothing factor (Γ_2^2 = noise smoothing in the beam after interception) predicted by Beam's theory have been plotted on the graph for each of the intercepting electrodes. The measured values of Γ_2^2 are in qualitative agreement with the functional variation predicted by the theory of Beam for magnetic field strengths greater than 150 gauss. The magnitude of the smoothing measured for the flat grid and for the .040-in. aperture lies within 1 db of the predicted value. The theoretical smoothing is consistently low in the case of the tilted grid, and high for the larger apertures. In the latter cases, modifying the theory to take into account the radial variation of current density has slightly improved the agreement between theory and experiment. The direct-current density in the actual beam was found to decrease much more slowly towards the beam edge than is predicted by theory. This may account in part for the discrepancy between the observed and predicted values of interception noise in cases where the beam is intercepted near the edge.

Concerning the magnitude of the interception noise produced by the mesh grids it is interesting to note that for field strengths in the neighborhood of 150 gauss the noise approaches or exceeds the North values plotted as curves A_2 and B_2 in Fig. 5. At these values of magnetic field, the theoretical mean spreading radius is nearly equal to the radius of one grid module. In the actual physical beam, small d-c. perturbations in the electron flow, due to imperfect electrostatic focussing fields in the electron gun, probably increase the spreading radius. Hence it would be expected from theoretical arguments that the noise should approach the North value.

The increase in the mean spreading radius due to transverse velocity other than that of thermal origin is believed to account for the fact that, for strong magnetic fields, the experimental values of Γ_2^2 for the grids are higher than the theoretical values.

From Fig. 5 it may be seen that the experimental values of smoothing for each of the intercepting electrodes decreased as the strength of the magnetic field decreased from about 150 gauss to zero. This is a region of transition from a mean spreading radius determined by the magnetic field to a mean spreading radius determined from transit time across the gun. The values of smoothing predicted from transit time considerations are in qualitative agree-

ment with the experimental measurements but are in error by 1.5 db for the tilted grid and the .040-in. aperture.

(*b*) *The Variation of Interception Noise with Gun Voltage*

The significant conditions of the experiment were: interception plate and cavity at the anode of the tantalum-cathode gun, magnetic field constant at 700 gauss. Figure 6 shows the collected results of smoothing measurements,

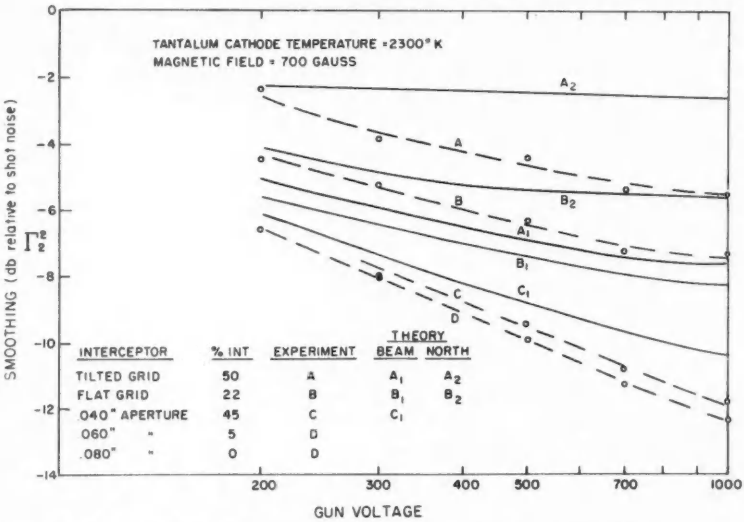


FIG. 6. The variation with gun voltage of noise smoothing after interception.

similar to those indicated in Fig. 4, for gun voltages from 200 volts to 1000 volts. The small amount of excess noise caused by aperture interception is again apparent. For 5% interception by the .060-in. aperture no change in smoothing was detectable within the precision of the measurements, which was ± 0.2 db. The .040-in. aperture which intercepted 45% of the beam current caused less than 0.5-db decrease in smoothing.

The results of a comparison of the experimental values of Γ_2^2 with the values predicted by the theories of Beam and North are not significantly different from those of the previous section.

(*c*) *Space-Charge Waves on an Intercepted Beam*

(*i*) *Space-Charge-Limited Emission*

Using the tilted grid and the .040-in. aperture as intercepting electrodes, noise was measured as the electron gun was moved away from the cavity. This was done for two cases: first, with the intercepting electrode at, and moving with, the anode of the electron gun; second, with the intercepting electrode stationary at the cavity. Figure 7 shows the maxima and minima of the noise space-charge waves for a gun voltage of 700 volts. Ordinate values

are Γ_2^2 = noise power relative to full shot noise in the transmitted current. It is to be emphasized that this graphical representation of the levels of the maxima and minima does not show the complete form of the space-charge waves, nor the fact that the wavelength is different in each case.

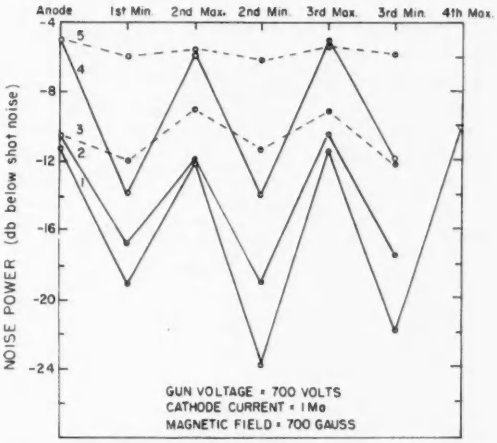


FIG. 7. Maxima and minima of the noise space-charge wave in the drift space for various conditions of interception.

Curve	Interceptor	Position	% Interception
1	.080-in. aperture	—	0
2	.040-in. aperture	At anode	45
3	.040-in. aperture	At cavity	45
4	Tilted grid	At anode	50
5	Tilted grid	At cavity	50

The .040-in. aperture, intercepting 50% of the beam current at the anode (curve 2), had only a small effect on the space-charge wave. The level of the minima was increased by approximately 2 to 4 db in 20 db. In the case of grid interception at the anode (curve 4), the over-all noise level was much higher, but the ratio of maximum power to minimum power was still large. The standing-wave ratio was in fact greater than that for the first minimum of the unintercepted beam (curve 1).

When the beam was intercepted at the cavity (curves 3 and 5) the standing-wave ratio was greatly reduced. With the tilted grid at the cavity, the standing-wave ratio was of the order of 0.5 db as compared with 10 db for the grid at the anode.

The gross effects of interception on the space-charge wave can be explained on the basis of the well-known theory of space-charge waves. At the intercepting electrode a new source of noise is created which establishes its own standing wave in the drift space. The total space-charge wave is the sum of the wave due to interception and the wave due to the noise in the beam before interception. The phase of the interception-noise wave is such that it

always has a maximum at the intercepting electrode. As long as the plane of interception occurs at a maximum of the standing wave in the unintercepted beam, the resultant standing wave still exhibits deep minima. When the interception takes place at the cavity, the noise due to interception is a maximum and almost completely masks the space-charge wave due to initial noise.

(*ii*) *Interception Noise on a Temperature-Limited Beam*

Measurements of the variation of noise power with collector current such as those illustrated by Fig. 4 showed that, for a given value of collector current in the temperature-limited region, the noise power measured at the anode was not increased by interception. That is, the noise at the anode is full shot noise. Beyond the anode there are minima of the space-charge wave where the noise is less than full shot noise.

The noise level at the first standing-wave minimum in a temperature-limited beam was measured for various conditions of interception. A constant collector current of 50 μa was used. The cathode current was larger than this depending on the particular intercepting electrode used. Results are listed in Table II.

TABLE II

Interception noise on a temperature-limited beam. Γ_2^2 is the noise level of the first standing-wave minimum in decibels below shot noise. Gun voltage = 700 volts; collector current = 50 μa; magnetic field = 700 gauss

Interceptor	Percentage interception	Γ_2^2 experimental	
		Interceptor at anode	Interceptor at cavity
.080-in. aperture	Zero	12.7	12.7
.060-in. ,,	15*	13.0	7.9*
.040-in. ,,	57	14.0	5.0
Tilted grid	45	11.5	3.0
Flat grid	20	12.2	5.5

*These values are approximate averages which take into account large fluctuations in intercepted current due to scalloping on the beam.

These measurements provide conclusive evidence that interception noise can arise from a smoothing process external to the electron gun. The reduction of noise at the minima of the space-charge wave results from interactions which are completely independent of processes in the electron gun.

The magnitude of the interception noise measured at the space-charge-wave minimum was considerably higher than that measured at the gun anode for space-charge-limited emission. Consideration of the behavior of a Pierce-type electron gun for low current densities has shown that there exists a highly distorted focussing field which produces excess transverse velocity. The consequent large increase in the mean spreading radius accounts for the higher values of interception noise. The experimental values of smoothing for the two grids are within 0.4 db of the North-theory value, which is indicative of the fact that the mean spreading radius was much larger than the effective radius of one grid module.

VII. CONCLUSIONS

The interception by a circular aperture of a fraction of the current in an electron stream causes little increase in noise when the electron beam is confined by a strong magnetic field. Fifty per cent interception produced a decrease of only 0.5 db in approximately 10 db of smoothing. A mesh grid caused decreases in smoothing as great as 7 db for an equivalent fraction of current intercepted.

The theories of North and Beam as modified by the author form a satisfactory basis for the prediction of the amount of interception noise caused by circular apertures and mesh grids. Perturbations in the electron flow, in particular the radial variation of direct current density, and excess transverse velocity due to imperfections in focussing fields have considerable effect on interception noise.

Interception of current excites a standing wave of noise, along a drifting electron stream, that has a maximum of noise power at the plane of interception. Thus, in beam-type microwave tubes, interception at the electron gun anode may be of little consequence, whereas interception by electromagnetic structures such as a resonant cavity may seriously degrade the noise performance of the tube.

VIII. ACKNOWLEDGMENTS

The writer wishes to thank Professor G. A. Woonton for his encouragement and interest during the work.

This research formed part of the work on the physics of electron beams, which is financed at the Eaton Electronics Research Laboratory by the Defence Research Board of Canada. The author is indebted to the National Research Council of Canada for granting him leave of absence in order to carry out this work, for financial assistance in the form of three studentships, and for permission to publish this paper.

REFERENCES

BEAM, W. R. 1955. RCA Rev. **16**, 551.
CUTLER, C. C. and QUATE, C. F. 1950. Phys. Rev. **80**, 875.
FRIED, C. and SMULLIN, L. D. 1954. I.R.E., Trans. on Electron Devices, ED-1, No. 4, 168.
KORNELSEN, E. V. 1957. Ph.D. Thesis, McGill University, Montreal, Que.
MCFARLANE, R. A. 1959. Ph.D. Thesis, McGill University, Montreal, Que.
MCINTOSH, B. A. 1958. Ph.D. Thesis, McGill University, Montreal, Que.
NORTH, D. O. *et al.* 1940. RCA Rev. **5**, 244.
PIERCE, J. R. 1949. Theory and design of electron beams (D. Van Nostrand Company, Inc., New York).
ROBINSON, F. N. H. 1954. J. Brit. Inst. Radio Engrs. **14**, 79.
ROBINSON, F. N. H. and KOMPFNER, R. 1951. Proc. I.R.E. **39**, 918.

FIRST ADSORBED LAYER OF HELIUM AT 4.2° K[1]

J. P. HOBSON

ABSTRACT

An attempt has been made to measure the helium adsorption isotherm on pyrex (Corning 7740) at 4.2° K for pressures above the adsorbed layer ranging from 10^{-4} to 10^{-12} mm Hg. New ultrahigh vacuum techniques were used for the pressure measurements with geometric adsorbing areas from 10 to 70 cm². Above 10^{-9} mm Hg the results give a true measurement. Below 10^{-9} mm Hg the results give an upper limit on the pressure above the adsorbed layer. The form of the isotherm indicates that the first adsorbed layer is complete at a pressure $\leqslant 10^{-9}$ mm Hg. A comparison of the present results with previous measurements of the helium adsorption isotherm at 4.2° K, which extend down to 10^{-6} mm Hg, suggests that all such measurements with solid adsorbents have been made on helium layers thicker than one monolayer. This conclusion supports similar conclusions drawn from heats of adsorption data at higher temperatures. An extrapolation of the results suggests that total gaseous densities approaching those of interstellar space (1.0×10^{-24} g/cm³) should be possible in the laboratory, with liquid helium trapping.

I. INTRODUCTION

Adsorption isotherms of helium have been measured by many authors for various adsorbents (Keesom and Schmidt 1933; Stout and Giaque 1938; ván Itterbeek and van Dingenen 1938; van Itterbeek, van Dingenen, and Borghs 1939; Keesom and Schweers 1941*a*, *b*; Kistemaker 1947; Long and Meyer 1949; Schaeffer, Wendell, and Smith 1949; Frederikse and Gorter 1950; White, Chien, and Johnston 1952; Bowers 1953; Brewer and Mendelssohn 1953; Tjerkstra, Hooftman, and van den Meydenburg 1953; Aston, Mastrangelo, and Tykodi 1955; Strauss 1956; Meyer 1956; Hoffman, Edeskuty, and Hammel 1956). The measurements taken together have covered a pressure range from near-atmospheric pressure to 10^{-6} mm Hg and a temperature range from 20° K to 1.2° K. Yet until recently the application of the available theories of adsorption left two basic questions unsatisfactorily answered:

(*a*) What is the density of atoms in the first adsorbed layer?

(*b*) At what pressure is this layer first complete?

These two questions were not answered satisfactorily even above the λ-point in a region where simple theories of adsorption might be expected to apply. The questions are related but (*a*) has been examined more extensively in the literature because the application of the well-known B.E.T. theory (Brunauer, Emmett, and Teller 1938) to the data yielded a first layer helium density that was anomalously high (Long and Meyer 1949; Band 1949; Aston and Mastrangelo 1951; Long and Meyer 1953; Atkins 1954; Singh and Band 1955).

Recently Steele (1956) and Meyer (1956) have independently proposed that the heat of adsorption of the second layer, although an order of magnitude

[1]Manuscript received November 10, 1958.
Contribution from the Electron Physics Group, Radio and Electrical Engineering Division, National Research Council, Ottawa, Canada.
Issued as N.R.C. No. 5082.

smaller than that of the first layer, remains larger than the heat of vaporization, and as a consequence the standard B.E.T. theory has tended to give the density of two layers rather than one. The bulk anomaly in the first layer density is thus removed. A consequence of this new proposal is that the first layer should be complete at extremely low pressures. Meyer (1956) examining his low-pressure data for gold concluded that the first layer is complete at $P/P_0 \sim 10^{-7}$ at 2° K and at $P/P_0 > 10^{-7}$ at 4.2° K. Steele (1956) does not treat (*b*) explicitly, but the question may be answered theoretically for TiO_2 at 4.2° K by combining the isotherm discussed by Steele with the force constants obtained experimentally by Aston, Mastrangelo, and Tykodi (1955). The same answer is obtained by using these constants with the standard B.E.T. isotherm or the Langmuir isotherm (1918), but because of the greater generality of the isotherm examined by Steele we will use it here. This isotherm equation is:

$$\frac{v}{v_m} = \frac{\left\{1+zx\left[\frac{2-x}{(1-x)^2}\right]\right\} cx}{1+cx+\frac{zcx^2}{1-x}} \tag{1}$$

where v is the quantity adsorbed, v_m is the quantity adsorbed in a monolayer, $x = P/P_0 =$ the saturation (P is the equilibrium pressure above the adsorbed layer, P_0 is the equilibrium pressure above the bulk liquid at the same temperature. $P_0 = 760$ mm Hg for helium at 4.2° K), z is a constant connected with the energy of adsorption of the second layer, c is the usual B.E.T. constant connected with the energy of adsorption of the first layer,

$$c = \exp\frac{(E_1 - E_{liq})}{RT} \tag{2}$$

where E_1 is the heat of adsorption of the first layer, E_{liq} is the heat of vaporization, R is the universal gas constant, and T is the absolute temperature. Aston, Mastrangelo, and Tykodi (1955) find E_1 to be 700–200 cal/mole. We have arbitrarily assumed $E_1 = 300$ cal/mole for calculation. From (2) this gives $c = e^{35} = 10^{15}$ at 4.2° K. Taking $v_m = .33$ cm³ S.T.P./m² and $z = e^4$ as given by Steele (1956) we plot equation (1), in Fig. 1. The knee at

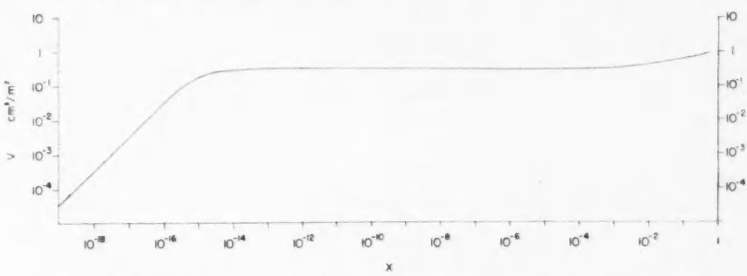


FIG. 1. Calculated helium adsorption isotherm for TiO_2 at 4.2° K. Curve gives equation (1) with $v_m = .33$ cm³/m², $c = e^{35} = 10^{15}$, $z = e^4 = 55$.

$x \approx 10^{-15}$ (i.e. $1/c$) represents the completion of the first layer. A non-constant E_1 as may be expected in practice will make the knee less clear but will not change its location, which lies well below the pressure range which has been studied by the authors quoted above.

The measurements of Hobson and Redhead (1958) on the operation of an inverted-magnetron pressure gauge involved adsorption of helium on pyrex at 4.2° K at saturations near 10^{-15}. However, in these experiments the quantity of helium available for adsorption was limited to that diffusing through the pyrex envelope (Rogers, Buritz, and Alpert 1954) and was too small to yield a monolayer on any measurable area. A simple modification of the apparatus for the introduction of larger amounts of helium promised at least an exploratory measurement of the helium adsorption isotherm on pyrex at 4.2° K for saturations from 10^{-7} to 10^{-15}, thus lowering the limit of previous measurements by six orders of magnitude, and providing an experimental check on the major feature of Fig. 1, namely the very small value of x at which the monolayer is complete. For physical adsorption this feature is expected to be approximately independent of adsorbent.

This paper describes such a measurement.

II. APPARATUS

The apparatus used is shown in Fig. 2. The envelope was pyrex glass (Corning 7740). The valves were ultrahigh vacuum valves (Alpert 1953;

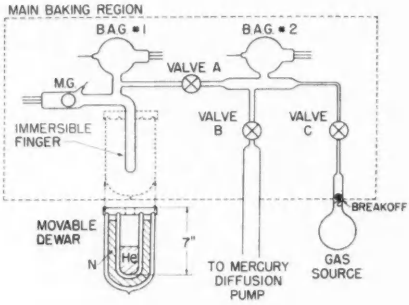


FIG. 2. Apparatus.

Bills and Allen 1955). The Bayard–Alpert gauges (B.A.G. No. 1, No. 2) were of standard design and were operated with collector at ground potential (i.e. 0), filament +45 v, grid +300 v, and grid current either 8 ma or 80 μa. The main adsorption volume V_1 (915 cm^3) was to the left of valve A and contained a magnetron pressure gauge (M.G.) (Redhead 1958), a B.A.G. (No. 1), and a finger which could be immersed in liquid helium at 4.2° K in a simple strip-silvered dewar.

V_1 was similar to the sealed-off systems used by Hobson and Redhead (1958) except that in the present experiments valve A permitted control of

helium flow into V_1. Valve C controlled flow from the gas source of spectroscopically pure helium to the central volume V_2 where pressure P_2 was measured by B.A.G. No. 2. V_2 could be pumped through valve B with a mercury diffusion pump suitably trapped with liquid nitrogen.

The system was prepared for measurements by baking the region shown in Fig. 2 for 15 hours at 400 to 450° C with all valves open and with the mercury diffusion pump running. Before measurements were begun, pressures of 5×10^{-11} mm Hg were obtained in V_1 with A closed, and 2×10^{-10} mm Hg in V_2 with C closed. It was found that diffusion of helium through the pyrex break-seal from the source flask at approximately atmospheric pressure was sufficient to raise P_2 to 10^{-7} mm Hg with C open, and to permit part of the experiment to be performed without breaking the seal.

In all liquid helium immersions the dewar position was fixed and the liquid level was permitted to fall by natural evaporation after transfer of about 170 cm^3 of liquid helium. A transfer lasted about one hour. The liquid level in contact with graduations on the finger could be observed visually immediately after transfer, but clouding in the dewar which was sealed at the top with a styrofoam cap and glass wool sometimes made visual observations unreliable later. A single loop of .002-in. tantalum wire was wrapped horizontally around the finger about 1 in. from the tip. Superconductivity was not obtained, probably because of unsuitable annealing conditions (Webber 1947), but with 30 ma direct current the resistance rose by 50% when the level fell below the wire, and a further 25% when the level left the tip of the finger. The latter event was also recognizable by a marked change in the rate of desorption of gas from the finger. By combining these data the helium level could be estimated at any time during a run. Attempts to maintain a fixed helium level by continuous raising of the dewar were not successful.

In this paper, pressures at the gauges are quoted as equivalent nitrogen pressures, and those above the adsorbed layer as true helium pressures, unless otherwise stated.

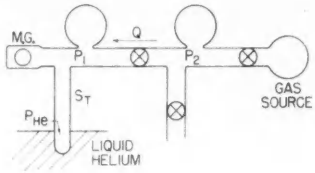
III. EXPERIMENTAL PROCEDURE

(1) *Calibration of Magnetron Gauge at Very Low Pressures*

In the main adsorption measurements, the M.G. was required to measure pressures below those measurable with the B.A.G. (lower reliable limit of B.A.G. No. 1 was about 3×10^{-10} mm Hg because of uncertainties in its X-ray limit), and in this range a new calibration procedure was used which did not simply check the magnetron gauge ion current (I_{MG}) against P_1 given by B.A.G. No. 1, as was done at higher pressures.

Valve A was closed and the finger immersed in liquid helium, with B.A.G. No. 1 turned off to eliminate any possible contaminants for its hot filament. I_{MG} fell to a low value $I_0 \sim 2\times10^{-13}$ amp corresponding to an unknown pressure P_0 (see Section V(a)) which was steady except for noise (signal-to-noise ratio about 3:1). With P_2 set to a steady helium pressure of about 10^{-8} mm Hg, valve A was cracked open and thereafter left untouched to permit a steady helium flow Q to enter V_1. The relations believed to determine P_1 are

given in Fig. 3, which shows a simplified sketch of Fig. 2. Here k is the conductance of valve A, S_{MG} and S_T are the pumping speeds of the M.G. and the liquid helium trap respectively, and P_{He} is the pressure at the bottom of the finger. (The validity of the remark that $P_{He} \approx 0$ can be checked later from the adsorption results.) The values of I_{MG} and P_1 being discussed are values in excess of I_0 and P_0 and thus result from Q.



$$P_2 \gg P_1$$

$$Q = k(P_2 - P_1) \approx kP_2$$

$$P_1 = \frac{Q}{S_{MG} + S_T}$$

$$S_T \gg S_{MG}$$

$$P_1 = \frac{Q}{S_T} = \frac{kP_2}{S_T}$$

(SINCE $P_{He} \approx 0$ S_T = CONDUCTANCE OF TUBING TO FINGER)

FIG. 3. Principle of gauge calibration at very low pressures. Calibration experiment: plot I_{MG} vs. P_2.

With A set, P_1 was proportional to P_2 and relative values of P_1 were obtained by adjusting P_2 with valve C. P_2 was readily measured with B.A.G. No. 2 and the major portion of the calibration of the M.G. consisted in plotted I_{MG} against P_2. As a general experimental check on the validity of the analysis, it was found that for at least 10 minutes after a change in P_2, I_{MG} remained time-independent, showing S_T to be sensibly independent of the liquid helium level. The value of the constant k/S_T could be obtained when P_1 was sufficiently high to be measured directly by B.A.G. No. 1.

The results of three such measurements are shown in Fig. 4. The values of P_2 shown are not exactly those observed, since the relation between I_{MG} and P_2 depended on the setting of valve A which was not exactly reproduced in each run. However, in each run $P_2 \approx 10^3\ P_1$ as shown. It can be seen that the near-linear law between I_{MG} and P_1 breaks down at a current of 1.1×10^{-9} amp, corresponding to a pressure of 4×10^{-10} mm Hg.

Sudden changes in the operating characteristics of magnetron-type gauges have been observed (Conn and Daglish 1954), but in the present case the question arose whether the break in Fig. 4 might possibly result from the calibration method itself, which was new. The method was applied to B.A.G. No. 1 and no break was found down to $P_1 \approx 1 \times 10^{-10}$ mm Hg. The M.G. and B.A.G. No. 1 were tested simultaneously and the break was present in the M.G. characteristic but not in the B.A.G. characteristic, although the

angle of 59° in Fig. 4 shifted to 54°. Next, the experiment was repeated with the M.G. operating with 5 kv between plate and cathode, but with 2060 gauss rather than 1060 gauss. The resulting calibration curve was quite different

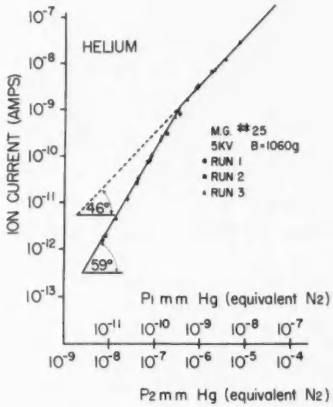


FIG. 4. Result of magnetron gauge calibration.

from that of Fig. 4. None of these data suggested any inherent error in the method of calibration, but all indicated that the calibration result was specific to M.G. operating conditions. The method of calibration was adopted as valid, and the low-pressure adsorption measurements were made with the M.G. operating conditions the same as those used during the calibration runs of Fig. 4.

(2) Adsorption Measurements

The objective of these measurements was to establish the equilibrium pressure above a layer of helium of known density per unit area, adsorbed on the inside of the pyrex finger.

Large adsorbed quantities were prepared by adjusting P_2 to a steady calculated value (B.A.G. No. 2 pumping very slowly with grid current 80 μa) with A, B, C closed. The finger was then immersed in liquid helium, producing conditions similar to those at the start of a calibration run. Valve A was next opened and the gas in V_2 allowed to be adsorbed in the finger. Valve A was then closed. After a few minutes the gauge current became constant except for noise and a slow increase caused by the falling liquid level. This pressure reading was converted to that above the adsorbed layer by corrections given below. When the pressure began to rise rapidly, indicating large desorption of gas from the finger as a result of disappearing liquid, the M.G., if operating, was switched off to prevent ionic pumping of the desorbing helium. When all liquid helium was gone, the pressure was measured by B.A.G. No. 1 operating at 80 μa. From this pressure reading, after a small correction for finger temperature, the quantity of gas adsorbed at any time could be deduced. This

pressure reading could be predictably calculated from the initial value of P_2 and the known volume ratio V_1/V_2, demonstrating that no helium was adsorbed irreversibly. This result was obtained previously by Hobson and Redhead (1958) for very small adsorbed quantities.

For small adsorbed amounts, P_2 was set to a fairly high value ($\sim 10^{-5}$ mm Hg) before A was opened partially for a short time and then closed. Experience permitted the quantity of gas so admitted to be judged quite accurately. This procedure was adopted for small quantities to maintain the ratio of helium to possible background gases in V_2 at a high value.

Certain assumptions concerning the pressure measurements and adsorbing areas were made which reduce the absolute accuracy of the results but which should yield valid relative results.

(*a*) *Pressure Measurements.*—Pressure measurements above the adsorbed layer were made with the M.G. alone at pressures below 5×10^{-10} mm Hg and with either or both gauges between this level and 10^{-7} mm Hg. Above 10^{-7} mm Hg errors might be expected from the high pumping speed of the M.G. ($\sim$.13 liter/second for helium) so only B.A.G. No. 1 was used. Since the gauges measuring P_1 were at room temperature (T_1 = 295° K) and the desired pressures P were those at T = 4.2° K, a thermomolecular pressure correction (Keesom 1942) was made according to the formula

$$\frac{P}{P_1} = \sqrt{\frac{T}{T_1}} = \sqrt{\frac{4.2}{295.}} = .119. \tag{3}$$

The calibration of the B.A.G. design used, as given by the manufacturer in equivalent nitrogen pressure, was adopted. However, ionization gauges are less sensitive to helium than to nitrogen and for the same ionization current the helium pressure is about 1/.16 times higher than the nitrogen pressure (Alpert 1953). This figure was adopted to convert observed ion currents to true helium pressures in the results presented below.

(*b*) *Adsorbing Surface Area.*—The adsorbing area has been taken as the simple geometric area of the inside surface of the finger (I.D. 2.2 cm, O.D. 2.5 cm) lying below the plane of the liquid helium surface. Adsorbing areas ranging from 10 to 70 cm^2 were used in calculating the results below.

IV. RESULTS

The results of 12 runs of the type described above are given in Fig. 5 where the number of atoms adsorbed per cm^2 is plotted against the measured pressure above the adsorbed layer, on log–log paper. The dotted line has been drawn visually. Three regions are apparent.

(*a*) At a pressure $P = 1.5\times10^{-12}$ mm Hg the points lie almost vertically and there is very little scatter.

(*b*) At pressures between 1.5×10^{-12} mm Hg and 5×10^{-10} mm Hg, the points lie approximately on a line at 45° to the horizontal and there is considerable scatter. The type of observation giving rise to this scatter is shown in Fig. 6, which compares two typical records of I_{MG} for approximately the same pressure: first, when the M.G. was measuring pressure with the finger immersed

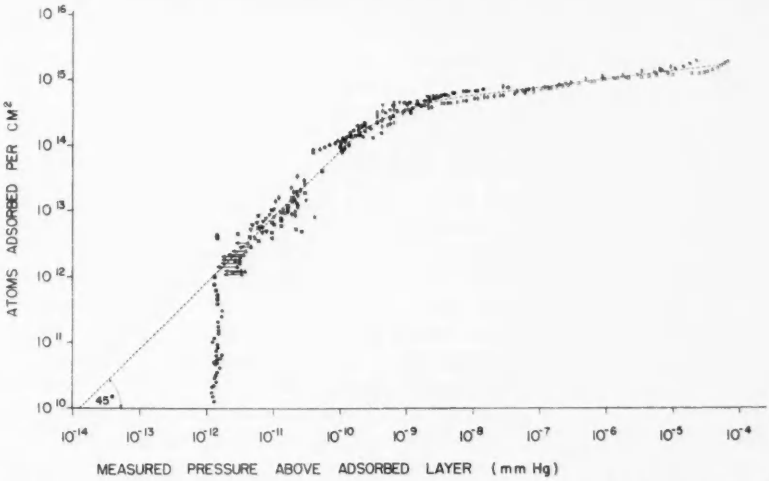


FIG. 5. Adsorption isotherm measurements for helium on pyrex at 4.2° K. Each symbol represents a different run. Solid symbols were obtained with magnetron gauge. Open symbols were obtained with Bayard–Alpert gauge.

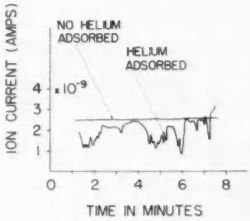


FIG. 6. Record of fluctuations in pressure above the adsorbed layer.

in liquid helium, and second, when there was no helium adsorption. Points in Fig. 5 were obtained from those in Fig. 6 by taking pressure readings at 1-minute intervals.

(*c*) Above 5×10^{-9} mm Hg the points lie approximately on a line at 7° to the horizontal and again there is comparatively little scatter.

V. INTERPRETATION OF RESULTS

The proposed interpretation of these three regions is as follows.

(*a*) The limiting pressure 1.5×10^{-12} mm Hg represents an instrumental limit of the apparatus. An approximate calculation of the residual pressure to be expected when the helium diffusing through the pyrex walls of V_1 flowed through the conductance of the finger gave an expected limit in Fig. 5 of 9.5×10^{-13} mm Hg. In this calculation the finger (2.2 cm I.D.) was assumed to have, in series, 9 cm at 295° K, 5.5 cm at 190° K, and 5 cm at 40° K, and

to be terminated by a plane at 4° K. The calculation assumes that the equilibrium pressure above the adsorbed layer may be neglected (Section V(*b*)). The resulting trap pumping speed was 4.8 liters/second. The measured helium diffusion rate was 4×10^{-10} mm Hg (equiv. N_2)/minute. Such high values have been discussed previously by Hobson and Redhead (1958). While the quantitative agreement between the observed and calculated instrumental limits is not fully satisfactory, nevertheless it does show that an instrumental limit is to be expected at roughly the pressure at which such a limit is observed. If this instrumental limit is taken as 1.5×10^{-12} mm Hg, then this pressure is to be subtracted from all other measured pressures above the adsorbed layer. The arrows in Fig. 5 show how a few points are shifted when this is done. The points are shifted progressively less for higher pressures.

(*b*) In this region between a pressure of 1.5×10^{-12} mm Hg and 5×10^{-10} mm Hg the amount adsorbed is approximately proportional to the measured pressure. If the heat of adsorption is independent of coverage, this result is to be expected for a partial monolayer (see Fig. 1). The absolute value of numbers of atoms adsorbed per cm^2 in this region is in general agreement with this assignment. This interpretation was at first placed on region (*b*) (Hobson 1958).

However, the following instrumental considerations place some doubt upon this interpretation. With an adsorbing area of 40 cm^2 corresponding roughly to the midpoint of a run, the adsorbing area was decreasing at about 2% per minute because of the falling liquid level. If the adsorbed layer is immobile (de Boer and Kruyer 1952), as seems likely from the considerations of Section VI, then at least $.02N$ atoms/minute are being released into the gaseous phase from a region near the liquid level, where N is the total number of adsorbed atoms.

Let us examine a model in which volume V_1 is imagined to be cut into two by the plane of the liquid level. Let the upper volume be U, which is bounded by the imaginary plane and all parts of V_1 not at 4.2° K ($U \sim 894$ cm^3). Let the lower volume be L, which is bounded by the imaginary plane and all parts of V_1 which are at 4.2° K. Further, let half the $.02N$ atoms released per minute be projected into U and the other half into L. Now assume that the pressure above the adsorbed layer is zero; i.e., that all atoms entering L never leave except by being released at the liquid level. U contains two pumps, namely the M.G. and the imaginary plane. The pumping speed of the M.G. is known to be about .1 liter/second, while the pumping speed of the plane of area A at $T = 4.2°$ K for helium ($M = 4$) is given by

$$F = 3.638A\sqrt{(T/M)} = 13.8 \text{ liters/second.} \tag{4}$$

Thus the total pumping speed of U is 13.9 liters/second. Consider now a point in region (*b*) of Fig. 5—say, 8×10^{13} atoms/cm^2 adsorbed at $P = 1\times10^{-10}$ mm Hg. The pressure P_1 as measured at the gauge is obtained from equation (3).

$$P_1 = 8.4\times10^{-10} \text{ mm Hg (true helium pressure).} \tag{5}$$

For an adsorbing area of 40 cm², $N = 3.2\times10^{15}$ atoms, giving a rate of input into U of 3.2×10^{13} atoms/minute or an equivalent room temperature leak rate of 1.7×10^{-8} mm Hg/second. The equilibrium pressure expected with a pumping speed given by equation (4) is

$$P_1 = [(.894)(1.7\times10^{-8})]/13.8 = 1.1\times10^{-9} \text{ mm Hg (true helium pressure).} \tag{6}$$

The general agreement between the observed value (5) and the value (6) calculated from a model in which the true pressure above the layer is negligible suggests that the pressures of region (*b*) be taken only as upper bounds on the true pressures above the adsorbed layer. The model which is linear in N predicts a 45° slope to the observations and provides a qualitative explanation for Fig. 6 since the gas causing the observed pressure originates at the liquid level where fluctuations are probable.

(*c*) In this region above 10^{-9} mm Hg the results represent true pressure measurements above the adsorbed layer.

VI. DISCUSSION

Figure 7 compares the present measurement with previous measurements on the physical adsorption of helium at 4.2° K. The scales are those commonly

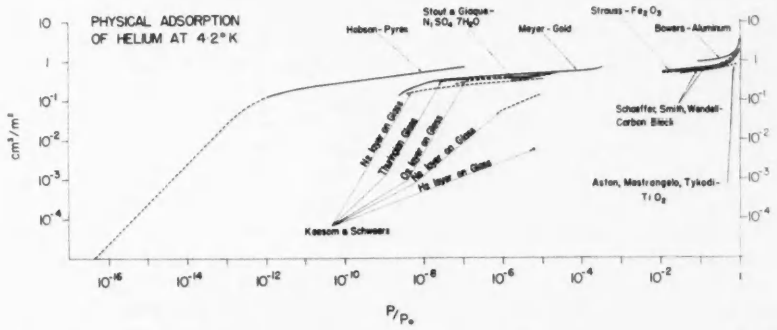


FIG. 7. Comparison of experimental results on the physical adsorption of helium at 4.2° K. The adsorbent material follows the authors' names. The dotted portion of the present results represents an upper bound on the pressures above the adsorbed layer. Van Itterbeek, van Dingenen, and Borghs (1939) and also White, Chien, and Johnston (1952) reported isotherms at 4.2° K on charcoal and silica gel respectively, but did not report the area of their adsorbent.

used in the adsorption literature. With the exception of Keesom and Schweers' (1941*b*) measurements with preadsorbed gases on thuringian glass, which are discussed below, the general similarity of Fig. 7 to Fig. 1 is plain. The essential parameter which determines the location of the knee in Fig. 1 is the magnitude of c which was obtained from the experimental heat of adsorption measured by Aston, Mastrangelo, and Tykodi (1955) for TiO_2 at 17° K. Thus the correlation between this heat of adsorption and the low temperature,

low pressure data appears to be experimentally established. It is not surprising therefore that some of the conclusions which are drawn below from an examination of Fig. 7, which gives results at one temperature only, have already been drawn by Aston and his colleagues from their measurements of heats of adsorption at much higher temperatures and pressures.

No changes in slope which could be interpreted as decisive evidence for the completion of the first adsorbed layer are seen, and it is proposed that no authors to date, probably including the present one, have made true pressure measurements above a partial monolayer of helium on a solid adsorbent for $T = 4.2°$ K. At temperatures below 4.2° K the first adsorbed layer is expected to be complete at even lower saturations than at 4.2° K. It may be noted that the results of Meyer (1956) on gold fit this conclusion without essential difficulty, although there is a suggestion that the first layer is complete at $P/P_0 \sim 10^{-9}$.

The present experiment gives the result that the first adsorbed layer is complete at $P/P_0 \leqslant 10^{-12}$, thus $c \geqslant 10^{12}$. From (2) this places a lower bound on the heat of adsorption in the first layer at 250 cal/mole. This is the highest heat of adsorption yet quoted from measurements in the temperature range $T \leqslant 4.2°$ K, the previous high being 148 cal/mole (Stout and Giaque 1938). However, heats of adsorption obtained near 17° K place the heat of adsorption between 700 and 200 cal/mole in the first layer (Aston, Mastrangelo, and Tykodi 1955; van Itterbeek and van Dingenen 1938). An evaluation of the heat of adsorption on charcoal by Steele and Halsey (1954) from apparent volume measurements gives 600 cal/mole. These results support the conclusion that only at temperatures well above 4.2° K have pressure measurements been successfully made above partial monolayers of helium.

From the results of Keesom and Schweers (1941*b*) shown in Fig. 7 it is concluded that Ne and H_2 preadsorbed on thuringian glass reduce the heat of adsorption for the first helium layer so that this layer is completed at $P/P_0 > 10^{-5}$, while with preadsorbed O_2 the first layer is complete at $P/P_0 < 10^{-7}$, and with preadsorbed N_2 at $P/P_0 < 10^{-9}$. It is interesting to present here Table I which gives the relative permeabilities through glass and the molecular diameters of these four gases and helium. Table I is taken from a paper by Norton (1957). The interpretation usually placed on the correlation in Table I is that diffusion takes place through holes in glass which are large enough to

TABLE I

Permeation of gases through silica glass

Gas	Permeation K through silica glass at 700° C	Atomic or molecular diameter, Å
Helium	2.1×10^{-8}	1.95
Hydrogen	2.1×10^{-9}	2.5
Neon	4.2×10^{-10}	2.4
Oxygen	Under 10^{-15}	3.2
Nitrogen	Under 10^{-15}	3.4

K = cm³ gas (S.T.P.) per second per cm² area per mm thickness

pass the smaller molecules but not the larger. On the surface of the glass these holes appear as sites (Tilton 1957) upon which physical adsorption might be expected to occur preferentially, and it is interesting to find that preadsorbed H_2, Ne, O_2, N_2 apparently lower the heat of adsorption of helium on thuringian glass in the same order that they are arranged in Table I, suggesting saturation of the adsorption sites by the preadsorbed gas. Steele and Aston (1957) have examined the effect on the heat of adsorption of helium on TiO_2 caused by preadsorption of argon, and have arrived at similar conclusions.

It is desirable to obtain an accurate measure of v_m from adsorption isotherms in the vicinity of the completion of the first layer, both because v_m is an important quantity, and in order to check Steele's (1956) values of v_m obtained at higher saturations. The present measurements are quite inadequate for the accuracy required because (*a*) all measurements were pressure measurements which were not calibrated directly against an absolute manometer, (*b*) the adsorbing area was not known exactly, (*c*) the scatter in the observed results was serious, (*d*) it is possible that the first layer density was determined by the density of adsorption sites rather than by interaction of adsorbed helium atoms.

It is now clear that accurate measurements of v_m for $T \leqslant 4.2°$ K in the vicinity of monolayer completion will require accurate pressure measurements below 10^{-9} mm Hg.

If the general interpretation placed on the results in Section V is correct, then an upper limit may be placed on the equilibrium pressure achievable above the minimum quantity adsorbed in Fig. 5. It is obtained by drawing a straight line horizontally from the point representing the minimum adsorbed density ($\sim 10^{10}$ atoms/cm^2) to the left until an intersection is made with the line at 45°. The pressure is about 10^{-14} mm Hg. This minimum quantity was determined by the helium diffusion into the system which was about an order higher than is usual for pyrex systems (Hobson and Redhead 1958). Further the normal diffusion rate may be reduced greatly by cooling the pyrex (Rogers, Buritz, and Alpert 1954) or by the use of materials other than pyrex for the vessel walls. With care the sources of other gases may be made small compared with the helium source even in a pyrex system, and these other gases may be expected to be pumped by liquid helium traps more effectively than helium itself. Thus the low pressure limitations of the present system regarding helium do not appear to be fundamental.

Gomer and his co-workers (e.g., Gomer, Wortman, and Lundy 1957) have used field emission microscopes with pyrex bulbs immersed in liquid helium. It is of interest to apply the results of Fig. 5 to their case. Their bulbs were approximately 100 cm^3 and spherical, and soon after seal-off could be cooled so that no further helium diffusion from the atmosphere took place. If it is assumed that the time between seal-off and cooling was 5 minutes, and that the helium leak rate was 10^{-10} mm Hg/minute (true helium pressure), then a helium pressure of 5×10^{-10} mm Hg was present at the time of cooling. When these atoms were absorbed there were about 2×10^7 atoms/cm^2 on the inside

surface of the bulb. By extrapolating Fig. 5 an upper limit on the helium pressure present is 2×10^{-17} mm Hg. This corresponds to a gaseous density of 5×10^{-24} g/cm^3, which may be compared with the mean density for interstellar gas of 1×10^{-24} g/cm^3 quoted by Allen (1955). Thus pressures in the laboratory approaching those of interstellar space do not appear unreasonable.

ACKNOWLEDGMENTS

The author wishes to thank Dr. D. Manchester, formerly of the National Research Council laboratories, for the initial stimulation of this work. Thanks are also due to P. A. Redhead, who designed the magnetron gauge, and E. V. Kornelsen, who assisted in some of the measurements. The apparatus was ably constructed by A. W. Pye and R. D. Cottee. Liquid helium was kindly supplied by the Low Temperature Group of the Pure Physics Division of these laboratories.

REFERENCES

ALLEN, C. W. 1955. Astrophysical quantities (Athlone Press, University of London) p. 227.
ALPERT, D. 1953. J. Appl. Phys. **24**, 860.
ASTON, J. G. and MASTRANGELO, S. V. R. 1951. J. Chem. Phys. **19**, 1067.
ASTON, J. G., MASTRANGELO, S. V. R., and TYKODI, R. J. 1955. J. Chem. Phys. **23**, 1633.
ATKINS, K. R. 1954. Can. J. Phys. **32**, 347.
BAND, W. 1949. Phys. Rev. **76**, 441.
BILLS, D. G. and ALLEN, F. G. 1955. Rev. Sci. Instr. **26**, 654.
BOWERS, R. 1953. Phil. Mag. **44**, 485.
BREWER, D. F. and MENDELSSOHN, K. 1953. Phil. Mag. **44**, 340.
BRUNAUER, S., EMMETT, P. H., and TELLER, E. 1938. J. Am. Chem. Soc. **60**, 309.
CONN, G. K. T. and DAGLISH, H. N. 1954. Vacuum, **4**, 136.
DE BOER, J. H. and KRUYER, S. 1952. Proc. Koninkl. Ned. Akad. Wetenschap. B, **55**, 451.
FREDERIKSE, H. P. R. and GORTER, C. J. 1950. Physica, **16**, 402.
GOMER, R., WORTMAN, R., and LUNDY, R. 1957. J. Chem. Phys. **26**, 1147.
HOBSON, J. P. 1958. Bull. Am. Phys. Soc. **3**, 339.
HOBSON, J. P. and REDHEAD, P. A. 1958. Can. J. Phys. **36**, 271.
HOFFMAN, C. J., EDESKUTY, F. J., and HAMMEL, E. F. 1956. J. Chem. Phys. **24**, 124.
KEESOM, W. H. and SCHMIDT, G. 1933. Proc. Koninkl. Ned. Akad. Wetenschap. **36**, 832.
KEESOM, W. H. and SCHWEERS, J. 1941*a*. Physica, **8**, 1020.
——— 1941*b*. Physica, **8**, 1032.
KEESOM, W. H. 1942. Helium (Elsevier Pub. Co., Inc., Amsterdam) p. 124.
KISTEMAKER, J. 1947. Physica, **13**, 8.
LANGMUIR, I. 1918. J. Am. Chem. Soc. **40**, 1361.
LONG, E. A. and MEYER, L. 1949. Phys. Rev. **76**, 440.
LONG, E. and MEYER, L. 1953. Advances in Phys. **2**, 1.
MEYER, L. 1956. Phys. Rev. **103**, 1593.
NORTON, F. J. 1957. J. Appl. Phys. **28**, 34.
REDHEAD, P. A. 1958. Report on 18th Annual Conference on Physical Electronics, Massachusetts Institute of Technology, Boston, p. 135.
ROGERS, W. A., BURITZ, R. S., and ALPERT, D. 1954. J. Appl. Phys. **25**, 863.
SCHAEFFER, W. D., SMITH, W. R., and WENDELL, C. B. 1949. J. Am. Chem. Soc. **71**, 863.
SINGH, R. P. and BAND, W. 1955. J. Phys. Chem. **59**, 663.
STEELE, W. A. and HALSEY, G. D. 1954. J. Chem. Phys. **22**, 979.
STEELE, W. A. 1956. J. Chem. Phys. **25**, 819.
STEELE, W. A. and ASTON, J. G. 1957. J. Am. Chem. Soc. **79**, 2393.
STOUT, J. W. and GIAQUE, W. F. 1938. J. Am. Chem. Soc. **60**, 393.
STRAUSS, A. J. 1956. Thesis, Chicago. (See Long and Meyer, 1953.)
TILTON, L. W. 1957. J. Research Natl. Bur. Standards, **59**, 139.
TJERKSTRA, H. H., HOOFTMAN, F. J., and VAN DEN MEYDENBURG, C. J. N. 1953. Physica, **19**, 935.
VAN ITTERBEEK, A. and VAN DINGENEN, W. 1938. Physica, **5**, 529.
VAN ITTERBEEK, A., VAN DINGENEN, W., and BORGHS, J. 1939. Physica, **6**, 951.
WEBBER, R. T. 1947. Phys. Rev. **72**, 1241.
WHITE, D., CHIEN, C., and JOHNSTON, H. L. 1952. J. Chem. Phys. **20**, 1819.

PROLONGED SIGNAL FADE-OUT ON A SHORT MICROWAVE PATH[1]

D. R. HAY[2] AND G. E. POAPS

ABSTRACT

During a period of 1 year, the incidence of signal fade-out has been observed in 2000 Mc/s transmissions over a 21-mile path near Ottawa. Fade-out durations varied from a few minutes to several hours, with the most frequent occurrence in summer and during the night. The cause of signal fade-out has been studied qualitatively through a ray analysis of the air refractivity profiles at the center of the radio path. This analysis indicates that fade-out is associated with a shallow, horizontal transition zone in air vapor pressure at a level near the antenna heights. Here, the vapor pressure changes through 1 millibar or more in a height interval of 100 ft. The signal fade-out is weak if the transition is from dry air below to moist air above, but is strong if the inverse transition occurs. Weak fade-out (a 'radio hole') is due to ray divergence, and strong fade-out (a 'radio antihole') is due to ray interference at the receiver.

INTRODUCTION

The infrequent disappearance of signal in transmission over a line-of-sight path is an accepted characteristic of microwave propagation. Normally, the signal is relatively steady; variation in amplitude is less than a few decibels about the mean level. The signal does, however, decrease by many decibels below the mean level on some occasions, even when atmospheric absorption is negligible. At such times, the signal may remain weak for several hours before returning to the normal level. These depressions in signal transmission are called signal fade-out, when the level of the signal decreases by more than five decibels from the mean. The term 'prolonged signal fade-out' is applied to this characteristic of the signal to distinguish it from the more rapid fading which occurs continually in transmissions over much longer ranges.

The cause of prolonged signal fade-out has not been understood, although this characteristic of signal transmission has been observed for several years. During the period 1939–45, many records of signal fade-out having durations of a few minutes were obtained on overwater and overland paths (Kerr 1951, Chapter 4). Prolonged signal fade-out was reported at a later date by Bean (1954) in the United States, and by others in Australia and New Zealand (Voge 1953). It is the practical difficulty of examining the radio path at the appropriate time that has delayed an extensive study of the problem.

This paper presents a qualitative study of microwave signal fade-out as carried out in eastern Ontario. Here, particular attention will be paid to the physical properties of the air near the earth's surface which have favored this weakening in signal transmission.

[1]Manuscript received December 3, 1958.
Contribution from the Defence Research Telecommunications Establishment of the Defence Research Board, Ottawa. The work was carried out under Project D 48-28-01-10.
[2]Present address: Department of Physics, University of Western Ontario, London, Ontario.

STATISTICS ON FADE-OUT

The incidence of fade-out was studied in signal transmissions over a 20.7-mile path near Ottawa. For this purpose, a 2000 Mc/s signal was directed from Franktown towards Hazeldean between 8-ft paraboloidal antennas. These antennas were mounted over 200 ft above ground to ensure a clear intervening path. Recording of the received signal power was continued from October 1956 until October 1957.

Two distinctive types of signal fade-out were found in the microwave transmissions. Sample records that illustrate these two types are shown in Fig. 1. The normal signal is relatively steady at the zero-decibel level on each record. Figure 1(*a*) indicates a weak fade-out, in which the signal amplitude decreased as much as 20 db below the mean level. It will be noted that changes in signal amplitude are gradual. The second type, a strong signal fade-out, is

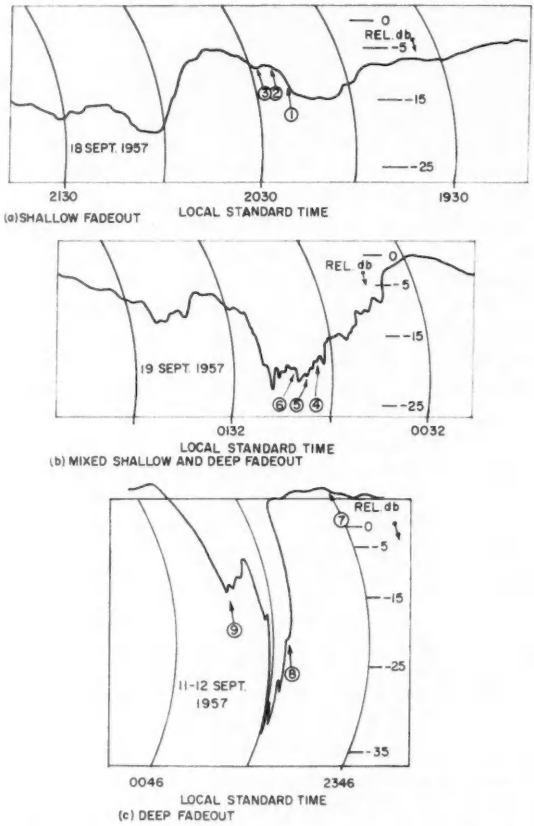


FIG. 1. Records of signal amplitude indicating types of fade-out.

shown in Fig. 1(*c*). Here the signal depression is more abrupt, and the minimum level is weaker than 30 db below the mean. The signal just before and after the fade-out is above the mean level. Figure 1(*b*) illustrates a signal fade-out that appears to be a combination of these two types.

The incidence of fade-out in the Ottawa area followed generally the pattern which has been reported for the central United States (Bean 1954). Figure 2 indicates the marked seasonal tendency. It will be seen that the signal was in fade-out for more than 24% of the time in summer, but for less than 5% of the time during the winter. The diurnal variation in fade-out incidence is shown in Fig. 3, where maximum frequency is indicated for the hours between midnight and 0600 hours. No fade-out has been observed on the Ottawa path between 1000 and 1600 hours. The onset of fade-out after sunset was spread over a longer interval than was the disappearance after sunrise.

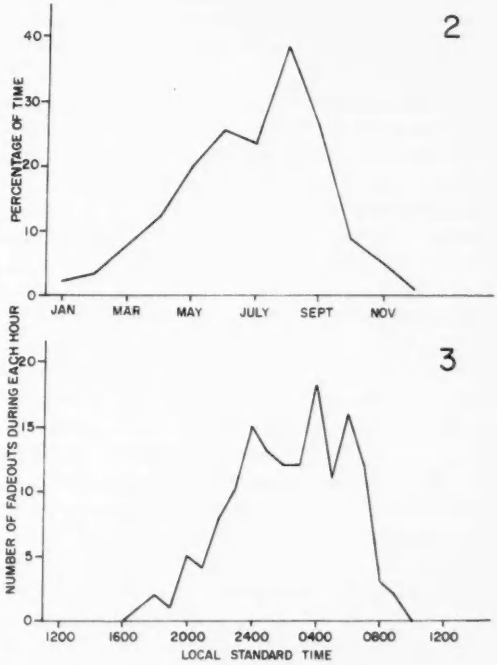


FIG. 2. Seasonal variation in signal fade-out (1957).
FIG. 3. Diurnal variation in incidence of signal fade-out (total for 1957).

Further analysis of fade-out on the Ottawa path has shown that the average duration was between 8 and 9 hours. The duration was generally independent of the maximum depth of the fade-out, and of the time of year. The weak type of fade-out in Fig. 1(*a*) was found throughout the year, but strong fade-out (Fig. 1(*c*)) was entirely absent in winter. The longest period during which

the signal depression exceeded 30 db below the mean level was 6 minutes; such depressions occurred for less than 0.3% of the time in summer.

The dependence of signal fade-out upon locality was apparent from a similar study on two separated radio paths. The second path was 31 miles in length, and was located about 25 miles north of North Bay. This experimental link was 200 miles from the Franktown–Hazeldean link. It was found that over a period of 1 year, the seasonal and diurnal trends in fade-out incidence were similar on the two paths. However, a detailed examination showed that there was little correlation between the incidence of individual fade-outs at Hazeldean and at North Bay.

REFRACTIVITY PROFILES

Special soundings of the temperature and humidity profiles through the air at the radio path were made during August and September of 1957. The site for these measurements was located midway between the transmitter and receiver. Several of these soundings were carried out while the microwave transmission was depressed by fade-out.

Conventional radiosonde equipment was modified for these special measurements. Two radiosondes were combined into one instrument so that together they would indicate air temperature and humidity continuously.[3] Although the slow response of the sensing elements in the radiosonde did not permit this type of instrument to indicate the fine structure in the air profiles, the radiosonde nevertheless was capable of sensing the principal features through transition layers in the air. The soundings at the center of the Franktown–Hazeldean path were made by controlled ascent and descent of the radiosonde, at a rate which permitted the sensing elements to approach equilibrium with the sampled air.

Some notable features were apparent in the temperature and dewpoint profiles of the air at the time of signal fade-out. In all, 27 soundings were made during August and September; seven of these were associated with prominent signal depressions at the Hazeldean receiver. Examination of these profiles showed that the signal transmission between Franktown and Hazeldean was normal when the air was well mixed and the lapse rates of temperature and humidity were approximately uniform through the lowest 1000 ft of the air. However, a prominent departure from the uniform lapse rates, at an altitude near the heights of the transmitting and receiving antennas, was observed whenever the microwave signal at Hazeldean was in fade-out.

Refractivity profiles have been derived from the temperature and dewpoint soundings, for a ray analysis of the radio path. The air pressure profiles, which are required also in computing refractivity, have been obtained from conventional radiosonde measurements at a nearby weather station. For the purpose of this ray analysis, curvature of the earth's surface and of the

[3]The pressure-operated switches in two Bendix–Friez radiosondes (model 279) were removed, and one radiosonde was set to measure temperature only while the other indicated only the humidity of the air. The height of the radiosonde was determined to an accuracy of 3% through the combined measurements of an optical inclinometer on the ground and a gauge that measured the length of the tethering cord (Anderson 1947).

layer of air above it is simulated by converting the refractivity profile of the air into a modified index profile (M-profile), and by taking the mean sea level as a flat surface (Kerr 1951, Section 2.4). The three M-profiles in Fig. 4 illustrate the main features of all of the profiles derived from the temperature and dewpoint soundings during signal fade-out. These profiles apply to the radio path at the times of the three signal recordings illustrated in Fig. 1.

RADIO RAY PATHS

The radio ray paths in the vertical plane between the Franktown transmitter and the Hazeldean receiver have been examined for an explanation of the observed signal fade-out. In this analysis it has been assumed that the physical structure of the air is spherically symmetrical and parallel to the mean sea level. Hence the vertical M-profile should be the same everywhere between the transmitter and receiver.

The diagrams on the right in Figs. 4(*a*), 4(*b*), and 4(*c*) represent the ray paths for the indicated M-profiles. Here, the rays have been constructed by the methods of geometrical optics (Kerr 1951, Section 2.3), since significant changes in refractivity take place over distances that are large compared with the radio wavelength.[4] It will be noted that the rays diverge from the transmitter with equal increments in angle of elevation and that the rays illustrated here occupy only a fraction of the antenna beamwidth. Then the signal field intensity is uniformly divided between the rays near the transmitter. Signal reflection from the ground has been omitted in these diagrams because the reflection coefficient may be assumed small at microwave frequencies (Bullington 1954).

Examination of the ray diagrams in Fig. 4 shows that the prominent irregularity in the M-profile near the antenna heights is the cause of a disturbance in an otherwise regular pattern. Two distinct effects of this irregularity upon the ray path are illustrated here. In the first example, Fig. 4(*a*) shows that a region of weak field intensity has been created at the receiver by the divergence of neighboring rays. The second effect is indicated in Fig. 4(*c*), in which the irregularity has caused a bundle of rays to be diverted so that ray interference occurs at the receiver antenna. Figure 4(*b*) illustrates a combination of these two patterns. Elsewhere, the region of weak field strength (Fig. 4(*a*)) is termed a 'radio hole', and the region of ray convergence in Figs. 4(*b*) and 4(*c*) a 'radio antihole' (Wong 1958).

The abrupt irregularity in the M-profile of Fig. 4(*c*) suggests the possibility of partial reflection of the incident microwave signal. Northover (1955) has found that elevated layer discontinuities in the air act as Fresnel reflecting surfaces at meter wavelengths; the reflection at centimeter wavelengths is generally weaker although still appreciable. Since partial reflections are ignored in the preceding ray analysis, they will be considered here as a possible

[4]An extensive analysis of ray paths through numerous refractivity profiles in the United States has been carried out for ground-to-air and air-to-air transmissions, with the use of an analogue computer. A description of the technique and of the results has been reported recently by Wong (1958).

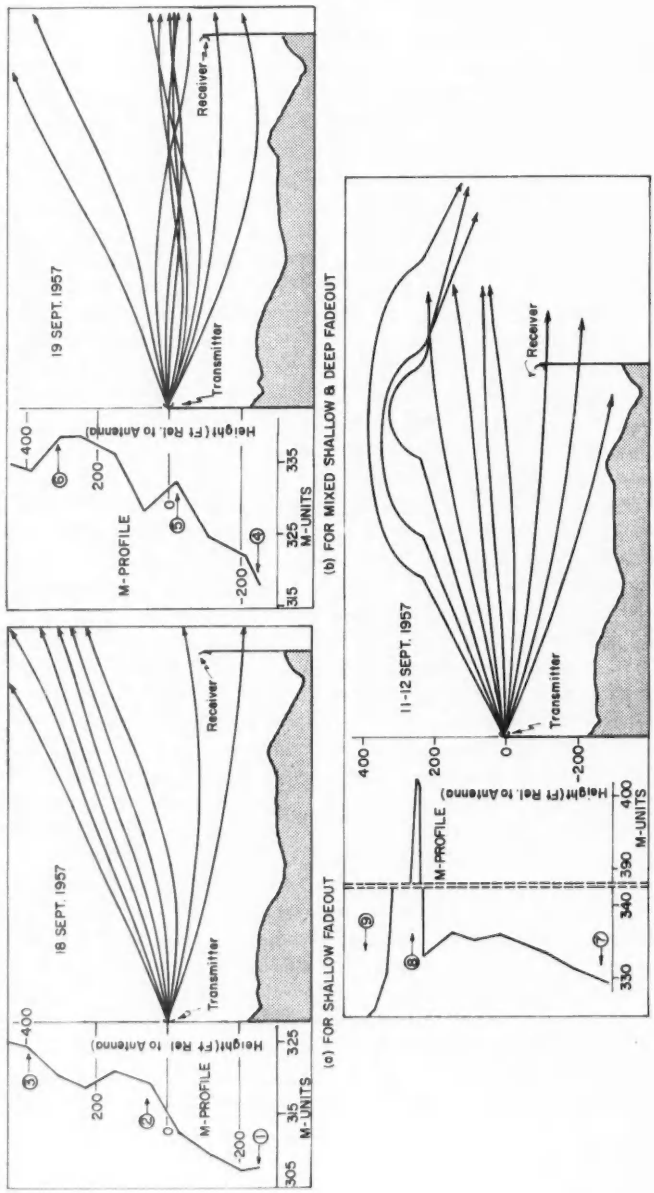


FIG. 4. *M*-Profiles and ray paths for signal fade-out.

cause of signal fade-out. The reflection coefficient of the discontinuity will be computed and specular reflection assumed to occur at the layer.

A simple calculation shows that partial reflection from an irregularity which is even more abrupt than that in the example of Fig. 4(*c*) cannot explain strong signal fade-out. If the *M*-profile in this figure contained the change from 333 to 400 *M*-units entirely at a height of 230 ft above the transmitter antenna, and if the refractivity is uniform above and below the boundary, the reflection coefficient of the boundary would be -0.43.[5] Interference between the reflected signal and the direct signal from the transmitter would cause a total change in the received signal of less than eight decibels. Indeed, a reflecting surface of this type would have to lie almost along the line joining the transmitter to the receiver in order to reflect sufficient signal for a strong fade-out.

From the above considerations, it appears that signal fade-out results from a disturbance of the radio transmission path by a prominent irregularity in the air refractivity. A ray analysis similar to those illustrated here has been carried out on all of the *M*-profiles associated with signal fade-out. The results indicate that weak signal fade-out (depression less than 20 db) is caused by simple weakening of the field at the receiver through ray divergence. Strong fade-out (depression greater than 30 db), however, requires the more complex multiple ray interference. The latter explanation is in accord with studies reported elsewhere on the synthesis of a deeply fading signal (Crawford and Jakes 1952; DeLange 1952; Kaylor 1953).

A note should be added here on the apparent inaccuracy of the ray diagrams. There appears to be no reasonable alternative to the explanation given above for signal fade-out, as based upon the ray analysis of the *M*-profiles. And yet, the receiver antenna generally is not in a region where fade-out would be predicted by the ray diagram. This departure must be attributed in part to the fact that the thin transition layers which lead to fade-out are rarely stationary, and to the difficulty in measuring temperature and dewpoint at moving layers with the modified radiosonde.[6] Inspection of the diagrams in Fig. 4 shows that the profile measurements at critical heights did not coincide with the time of maximum fade-out, and consequently the direction of the ray focusing would be altered during the sounding.

[5]From Stratton (1941), Section 9.5, the reflection coefficient (Γ) in which dielectric constant is replaced by refractive index (n) is

$$\Gamma = \frac{\sin\psi - \left[\left(\frac{n_2}{n_1}\right)^2 - \cos^2\psi\right]^{1/2}}{\sin\psi + \left[\left(\frac{n_2}{n_1}\right)^2 - \cos^2\psi\right]^{1/2}} = \frac{5\times10^{-3} - 1.27\times10^{-2}}{5\times10^{-3} + 1.27\times10^{-2}} = -0.43$$

where ψ = the grazing angle of reflection = 5 milliradians, and subscripts 1 and 2 refer to the media containing the reflected wave and the transmitted wave respectively. This equation applies to vertically or horizontally polarized waves in air.

[6]Soundings for each profile with the modified radiosonde required a total time of 10 to 45 minutes. The numbered arrows in Fig. 1 and at the *M*-profile of Fig. 4 indicate the position of the radiosonde during the signal recording interval.

CLIMATOLOGY

Information on the layer structure that leads to signal fade-out has been obtained from the radiosonde measurements. An inspection of the ray diagrams shows that changes in the amplitude of the radio signal at the receiver will be small unless a prominent departure from the mean occurs in the M-profile. It is necessary, further, that the departure appears near the line between the transmitter and receiver. Irregularities in refractivity at other heights have only a minor effect upon the ray diagram. The formation of a 'radio hole' (shallow fade-out in Fig. 4(a)) at the receiver accompanies a decrease in the slope of the M-profile over a shallow height interval, while a 'radio antihole' (deep fade-out in Figs. 4(b), 4(c)) is related to an inversion of the M-profile, also through a shallow height interval. All of the departures from mean refractivity that contributed to signal fade-out were less than 100 ft in thickness.

The temperature and humidity soundings indicate further that each of the significant departures in refractivity was due to a transition from dry to moist air through the shallow height interval. These transitions consisted of changes in the air vapor pressure of about one millibar or greater. Moist air is below the drier air for inversions in the M-profile, but drier air is below the moist air for a decrease in the slope of the M-profile. The example of Fig. 4(c) indicates a change in vapor pressure of 21 millibars at the lower boundary of the elevated irregularity, and a more gradual transition of the same order in the reverse direction throughout the rest of the irregularity.

A system of stationary air pressure over the transmission path is a general requirement for microwave signal fade-out. During the 12 months of the experiment, 64% of the fade-outs were simultaneous with a center of high pressure and 30% with the slowly changing transition interval between centers of high and low pressure. The remaining 6% followed shortly after a center of low pressure had moved through the radio path.

The motion of the shallow transition layers associated with signal fade-out should be commented upon. The radiosonde measurements of September 19 which yielded the M-profile in Fig. 4(b) also showed that the prominent irregularity near the height of the antenna appeared to descend through a vertical interval of 100 ft in 1 hour. During this time, only a minor change occurred in the refractivity profile through this irregularity. The apparent vertical motion may be due either to horizontal movement of an inclined layer or to the descending of a horizontal layer. In either case, the effect of the layer upon microwave signal transmission would be essentially the same, since the ray focusing is present even when the layer is inclined so that it is parallel to the mean slope of the land.

REMARKS

The ray analysis presented here indicates that microwave signal fade-out is associated with a shallow layer of transition in air vapor pressure. All of the observed transitions occurred through a height interval of the order of

100 ft; the change in vapor pressure was 1 millibar or greater. Further, these transitions were found within several hundred feet of the ground and near the height of the antennas. It is interesting to note that the condition for shallow signal fade-out, viz. a transition from dry air below to moist air above, is found at the latitude of Ottawa throughout the year. However, the radio transmission experiment suggests that the inverse transition for deep signal fade-out is absent during colder weather. Such shallow transitions in vapor pressure apparently depend upon vertical stability of the air, since they are formed generally at night in a system of stable air pressure, and they are dispersed rapidly as the sun heats the earth's surface in the morning.

ACKNOWLEDGMENT

The authors are grateful to Dr. W. Godson of the Meteorological Branch, D.O.T., for many valuable comments and suggestions.

REFERENCES

ANDERSON, L. J. 1947. Bull. Am. Meteorol. Soc. **28**, 356.
BEAN, B. R. 1954. Proc. Inst. Radio Engrs. **42**, 848.
BULLINGTON, K. 1954. Proc. Inst. Radio Engrs. **42**, 1258.
CRAWFORD, A. B. and JAKES, W. C. 1952. Bell System Tech. J. **31**, 68.
DELANGE, O. E. 1952. Bell System Tech. J. **31**, 91.
KAYLOR, R. L. 1953. Bell System Tech. J. **32**, 1187.
KERR, D. E. 1951. Propagation of short radio waves (McGraw-Hill Book Co., Inc., New York).
NORTHOVER, F. H. 1955. Can. J. Phys. **33**, 241.
STRATTON, J. A. 1941. Electromagnetic theory (McGraw-Hill Book Co., Inc., New York).
VOGE, J. 1953. Onde élect. **33**, 136.
WONG, M. S. 1958. Proc. Inst. Radio Engrs. **46**, 1628.

MULTIPLE NEUTRON CAPTURE IN THE MIKE FUSION EXPLOSION[1]

A. G. W. CAMERON

ABSTRACT

The constant ratio of yields of successive even or odd mass number nuclides produced in the Mike fusion explosion of 1952 implies an approximate constancy of the corresponding neutron capture cross sections for neutrons near 100 kev for the nuclei between U^{238} and U^{254}. A discussion is given of the physical quantities on which the neutron capture cross sections depend, particularly radiation widths, nuclear level spacings, and neutron binding energies. It is concluded that conventional mass formulas are probably wrong in predicting a large decrease in neutron binding energy as one goes from U^{238} to U^{254}. It is possible that the addition of a direct symmetry term to the mass formula might repair this difficulty. Another possibility is that an extra diffuseness of the nuclear surface in neutron-rich nuclei would reduce the Coulomb energy of the nucleus, thus stabilizing neutron binding energies.

YIELDS FROM THE MIKE EXPLOSION

In November, 1952, the United States Atomic Energy Commission conducted Operation Mike on an island in the Pacific Ocean. A refrigerated fusion device was exploded. The large neutron flux released in the explosion irradiated a mass of U^{238}. Many multiple neutron capture events occurred, forming heavy uranium isotopes which later decayed by beta emission, leading to the discovery of the elements Einsteinium and Fermium (Ghiorso, Thompson, Higgins, Seaborg, Studier, Fields, Fried, Diamond, Mech, Pyle, Huizenga, Hirsch, Manning, Browne, Smith, and Spence 1955). A partial list of isotope yields from this test has been deduced by the writer from the literature (Huizenga and Diamond 1957); these are plotted in Fig. 1. It may be seen

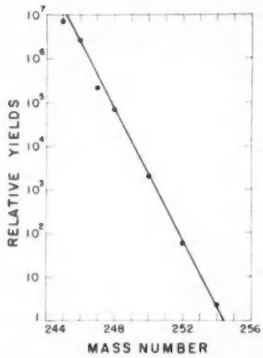


FIG. 1. Yields of nuclides of various mass numbers, produced in the Mike fusion explosion, deduced from abundances reported and interpolated by Huizenga and Diamond (1957).

[1]Manuscript received December 4, 1958.
Contribution from Atomic Energy of Canada Limited, Chalk River, Ontario.
Issued as A.E.C.L. No. 760.

that the yields of even mass numbers are an exponentially decreasing function of the mass number; the yields of odd mass numbers follow a similar relation but are depressed relative to the curve for even mass numbers by a factor 2. This reflects the usual odd–even variation of average neutron capture cross sections. Disregarding this odd–even effect, the yield drops by a factor of 5.8 per mass number. It has recently been revealed (Seaborg 1958) that the yields also follow this relation back to the vicinity of mass number 239.

The following is a speculative description of the operation of this fusion device. We may suppose that the basic essentials of the device consisted of a uranium or plutonium fission bomb, surrounded by a liquid deuterium–tritium mixture, surrounded in turn by a mantle of natural uranium. The explosion of the central fission trigger raised the central temperature into the range $5\times10^{7\circ}$ K to $10^{8\circ}$ K (Bethe 1958). The heavy uranium mantle held the D–T mixture in place while it was heated by various energy transfer mechanisms. Gryzinski (1958) has suggested that a nonthermonuclear fusion chain reaction will take place in a D–T mixture when the temperature exceeds $10^{7\circ}$ K. However, let us be conservative and neglect the possibility of this process taking place.

The D–T fusion reaction releases 3.6 Mev in the form of kinetic energy of the He^4 nucleus and produces a 14.7 Mev neutron. Again being conservative, let us assume that only the former amount of energy is available for heating the plasma. This energy is divided between kinetic energy of the particles and radiation energy density. At normal liquid hydrogen densities the plasma temperature can rise to $6\times10^{7\circ}$ K; if the D–T mixture should be compressed to 10 times its normal density the temperature can rise to $10^{8\circ}$ K. In the normal density case the mean reaction time will be about 1 microsecond (Post 1956), while in the compressed case it will be about 0.025 microsecond. It seems likely that some compression will have occurred since the D–T mixture was squeezed between the uranium mantle and the expanding gas from the fission explosion.

The energetic neutrons will have interacted for the most part with the nuclei of the mantle, producing fission and $(n, 2n)$ and $(n, 3n)$ reactions. The fission in the mantle produced most of the energy release in the explosion. However, for present considerations the important effect is that each 14-Mev neutron will have produced about three secondary neutrons mostly in the 1- to 2-Mev energy range. Inelastic scattering quickly reduced the energy of these neutrons to the vicinity of 100 kev (Vasilyev, Zamyartnin, Toropov, and Fomushkin 1957). Following several elastic collisions, these neutrons would be captured by the surrounding nuclei, provided they did not escape from the system. The U^{238} neutron capture cross section near 100 kev is about 0.2×10^{-24} cm^2 (Hughes and Schwartz 1957). Hence the 100-kev neutrons will have been captured after a mean life of about 0.25 microsecond in uranium of normal density. Since the uranium is likely to be compressed somewhat less than the D–T mixture, it may be seen that the mean D–T reaction time is comparable to the mean lifetime of the secondary neutrons. Hence the later stages of the capture will take place after the production of

14-Mev neutrons has nearly ceased. Since most of the production of the heaviest uranium isotopes by multiple neutron capture takes place in the very late stages of the capture process, it is reasonable to neglect any interference with this capture process by the declining flux of 14-Mev neutrons.

It appears that subsequent fusion explosions have not involved refrigerated devices, and it has been suggested that in these devices the deuterium–tritium mixture has been replaced by a Li^6–D mixture (or compound). The operation of this device would be similar to that described above except that the neutrons which are produced are necessary for the propagation of the explosion. These neutrons induce $Li^6(n, \alpha)H^3$ reactions, and the tritium so produced can then fuse with the deuterium. Because of this additional neutron destruction process, the neutron flux available for capture by heavy nuclei will be considerably reduced. As will be seen, multiple neutron capture yields are proportional to a high power of the neutron flux, and hence the yields will be very much reduced in this type of device. If these considerations are correct, it appears that the Mike explosion has been a unique and important scientific experiment.

If a mass consisting of some given species of nucleus is irradiated with a neutron flux ϕ for a time t which is small enough so that the depletion of the number of original nuclei can be neglected, then the following ratio is established between the abundances of the n and $n-1$ members of the multiple neutron capture product chain:

$$N_n/N_{n-1} = \sigma_{n-1}\phi t/n, \tag{1}$$

where σ_{n-1} is the appropriate average neutron capture cross section of the $n-1$ product nucleus. Since ϕt is the integrated neutron flux, it is apparent that this result remains unchanged if more complicated temporal variations of the neutron flux are considered.

If this were the situation in the Mike explosion, it may be noted that the product ratios could remain constant only if the average neutron capture cross sections near 100 kev increase by an order of magnitude as one goes from U^{238} to U^{254}. However, there are two modifying effects which must be considered.

The first effect is the modification introduced by exposing a mass of U^{238} to an integrated neutron flux which varies from one part to another. We may illustrate the effect of this by assuming an infinite mass of U^{238} exposed to a flux $\phi t = \alpha r^{-4}$ where α is a constant and the flux is cut off below $r = r_1$. Then the factor $\phi t/n$ of equation (1) is replaced by $\alpha(4n-7)/nr_1^4(4n-3)$. This is just the maximum integrated flux α/r_1^4 multiplied by the ratio $(4n-7)/(4n-3)n$. The latter ratio tends to $1/n$ at large values of n but varies much more rapidly for small values of n. If this were the only modifying effect we would require the cross sections to increase by an additional order of magnitude between U^{238} and U^{254} in order to give equal abundance ratios for successive capture products. Nearly all of this additional increase would be in the vicinity of U^{238}.

The second modifying effect is depletion of the source nuclei. From the above discussion we may see that the heaviest nuclei have been predominantly produced in the region of highest flux. Equation (1) may be crudely interpreted as indicating that the $n-1$ species has been present in the neutron flux for a time t/n. Neutron capture in this species has depleted it to a value $\exp(-\sigma_{n-1}\phi t/n) \approx 0.8$, where the numerical value is taken from Fig. 1 neglecting the odd–even effect. Hence $\sigma_{n-1}\phi t/n \approx 0.23$, for $n \sim 12$. On this basis the source nuclei in the region of highest flux have been depleted to a value $\sim \exp(-2.8\ \sigma_0/\sigma_{n-1})$. It may be seen that for $\sigma_0 \approx \sigma_{n-1}$ the source nuclei are depleted to about 6% of their initial abundance. The greatest depletions are in the vicinity of U^{238}. This effect compensates for the previously supposed necessity of increasing the capture cross sections in the heavier uranium isotopes.

A complete evaluation of these effects requires a knowledge of the integrated flux distribution in the Mike device. Even in the absence of this knowledge, however, some valuable conclusions can be drawn. The abundance ratio of equation (1) is equal to a cross section multiplied by a factor which can vary rapidly only near U^{238} due to any of the effects considered. Hence it is possible to conclude that the average capture cross section for 100-kev neutrons is nearly constant among the heavier uranium nuclei, say from U^{242} to U^{254}. This is very surprising, since, as the following discussion will show, one would expect the capture cross section to decrease steadily as the mass number of the uranium nucleus increases and its neutron binding energy decreases.

In a first attempt to understand this effect, based only on the data of Fig. 1, it may be supposed that there is a significant drop in neutron binding energy beyond U^{244}, which has 152 neutrons. There is a gap in the Nilsson diagram for deformed nuclei at this point (Mottleson and Nilsson 1958) which shows this effect among heavier nuclei on the valley of beta stability. If this were the case, it is conceivable that the neutron binding energy might be nearly constant for a considerable distance beyond U^{244}, leading to approximately constant neutron capture cross sections. But this explanation would require a sharp kink downwards in the yield curve as one goes beyond U^{244}, and this kink was not observed (Seaborg 1958). Hence it is very unlikely that there is an appreciable discontinuity of binding energy at U^{244}.

Since there is thus no known reason why the neutron binding energy and nuclear level density properties of the uranium nuclei should change abruptly immediately beyond U^{238}, it is a reasonable conclusion that the average neutron capture cross section at 100 kev is approximately constant from U^{238} to U^{254}. This will be assumed in the following discussion, but it should be noted that this discussion would not be changed in its essentials if the approximate constancy were restricted to the more certain range U^{242} to U^{254}.

The remaining part of the paper contains a discussion of the physics of neutron capture cross sections and a subsequent interpretation of the unexpected results from the Mike test.

NEUTRON CAPTURE CROSS SECTIONS

The neutron capture cross section averaged over resonances is (Lane and Lynn 1957):

$$\langle\sigma\rangle = \sum_l \sum_J \langle\sigma_{Jl}\rangle, \tag{2}$$

where

$$\langle\sigma_{Jl}\rangle = \frac{(2J+1)}{(2I+1)}\frac{\pi^2}{k^2}\frac{1}{\langle D_{\lambda J}\rangle}\sum_j \left\langle \frac{\Gamma_{\lambda J(nlj)}\Gamma_{\lambda J(\gamma)}}{\Gamma_{\lambda J}} \right\rangle, \tag{3}$$

and

$$\Gamma_{\lambda J} = \sum_{l,j} \Gamma_{\lambda J(nlj)} + \Gamma_{\lambda J(\gamma)}. \tag{4}$$

Here λ refers to the various resonances over which the average is taken, each of spin J, l is the angular momentum of the incoming neutron, k is 2π times the wave number of the incident neutron, I is the spin of the ground state of the target nucleus, $\Gamma_{\lambda J(\gamma)}$ is the radiation width of the resonance λ, $\Gamma_{\lambda J(nlj)}$ is the entrance channel neutron width for channel spin j (which may take the values $j = I \pm \frac{1}{2}$ except for the case $I = 0$ when $j = \frac{1}{2}$), $\Gamma_{\lambda J}$ is the total width of the resonance, and $\langle D_{\lambda J}\rangle$ is the average spacing between resonances of spin J and given parity at the excitation energy at which the capture is occurring. The contribution of inelastic scattering has been omitted from equation (4) on the grounds that we are interested in neutron capture processes after the neutron energies have been degraded by inelastic scattering to the point where the capture cross sections are comparable to or larger than the inelastic scattering cross sections. A contribution from fission has also been omitted since it seems unlikely that nuclei heavier than U^{238} have significant fission cross sections for 100-kev neutrons.

One commits only a relatively small error (Lane and Lynn 1957) if in equation (3) we make the replacement:

$$\sum_j \left\langle \frac{\Gamma_{\lambda J(nlj)}\Gamma_{\lambda J(\gamma)}}{\Gamma_{\lambda J}} \right\rangle \approx \sum_j \frac{\langle\Gamma_{\lambda J(nlj)}\rangle\langle\Gamma_{\lambda J(\gamma)}\rangle}{\langle\Gamma_{\lambda J}\rangle}. \tag{5}$$

In the vicinity of 100 kev in uranium nuclei, large contributions to the neutron capture come from s-, p-, and d-wave neutrons; the contribution from f- and higher-wave neutrons is negligible. This results from the fact that for s-, p-, and d-wave neutrons,

$$\langle\Gamma_{\lambda J(nlj)}\rangle \gg \langle\Gamma_{\lambda J(\gamma)}\rangle, \tag{6}$$

whereas for f-wave neutrons the smaller barrier penetrability reverses the inequality of relation (6). It follows that

$$\frac{\langle\Gamma_{\lambda J(nlj)}\rangle\langle\Gamma_{\lambda J(\gamma)}\rangle}{\langle\Gamma_{\lambda J}\rangle} \approx \langle\Gamma_{\lambda J(\gamma)}\rangle \tag{7}$$

for the lower neutron partial waves.

It has been found (Cameron 1957*a*, 1957*d*) that radiation widths are independent of J and of λ to a good approximation. Hence it follows from relations (2), (3), (4), (5), and (7) that

$$\langle\sigma\rangle \approx \text{const.}\,\frac{\langle\Gamma_\gamma\rangle}{\langle D_0\rangle}, \tag{8}$$

where for simplicity the radiation width has been written as $\langle\Gamma_\gamma\rangle$, and $\langle D_0\rangle$ is the average level spacing for the levels with $J = 0$. This ignores the fluctuation in the average of the statistical factor $(2J+1)^2$ which contributes to the odd–even variation in the capture cross sections ($\langle D_J\rangle \propto (2J+1)^{-1}$ according to Cameron (1958*a*)).

Therefore the results from the Mike test imply that $\langle\Gamma_\gamma\rangle/\langle D_0\rangle$ is approximately constant in the vicinity of 100 kev above the neutron separation energies for nuclei between U^{239} and U^{255}. In the following sections we will discuss some of the physical quantities associated with this ratio.

NUCLEAR LEVEL SPACINGS

It was shown by Newton (1956) that nuclear level spacings could be accurately calculated from the usual Fermi gas model if one used appropriate averages of the densities of actual nucleon orbits near the Fermi level. This procedure allows shell effects to be taken into account. The writer (Cameron 1958*a*) has used this formulation but has improved the values of the nucleon orbit spacings by a procedure which relates them to the masses of nuclei. The level spacing formula is

$$\langle D_J\rangle = \frac{3.26\times10^5}{(2J+1)}\left[G_Z G_N G^3 A^2\left(\frac{\pi^2}{3}Gt-\frac{3}{2}\right)t^7\right]\exp-\frac{\pi^2}{3}Gt, \tag{9}$$

where

$$G = G_Z+G_N, \tag{10}$$

and

$$t = \frac{3}{\pi^2 G}\left[\frac{3}{2}+\left(\frac{9}{4}+\frac{2\pi^2 GU}{3}\right)^{1/2}\right]. \tag{11}$$

Here G_Z and G_N are the average spacings of the proton and neutron orbits at the Fermi level of the nucleon gas, U is the effective excitation energy in Mev (defined as the actual excitation energy minus the pairing energy of the paired nucleons), t is the nuclear temperature, and $\langle D_J\rangle$ is in electron volts.

NUCLEAR RADIATION WIDTHS

In this section we consider at some length nuclear radiation widths. Since these play an important part in the further interpretation of the Mike yields, it is necessary to establish some confidence in the validity of the basic equation (12) below.

Nuclear radiation widths are slowly varying quantities. They have a general decreasing trend with increasing mass number, with superposed in-

creases just below closed neutron shells. The measured values collected by Levin and Hughes (1956) and by Stolovy and Harvey (1957) are plotted in Fig. 2.

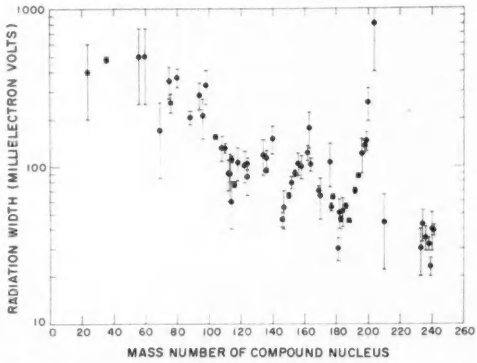


FIG. 2. Total radiation widths plotted from the compilations of Stolovy and Harvey (1957) and Levin and Hughes (1956).

The radiation process is rather complex, since there is usually a very large number of levels which are energetically accessible for radiative transitions from the capturing state. It has been shown that electric dipole transitions predominate in the capture process (Bartholomew, Campion, Knowles, and Manning 1957). In such transitions the angular momentum spin vector must change by one unit and the parity must change. Hence in general any given spin state radiates to states of opposite parity and three values of spin. In calculating total radiation widths it is necessary to sum over all these transitions.

Blatt and Weisskopf (1952) have argued that in general the matrix elements for the transitions to lower states will be reduced as the level spacing in the vicinity of the radiating state decreases. In the radiation process all but one of the nucleons form a core which remains unchanged; coupled to this core is the remaining nucleon which changes its angular momentum. The total radiation probability for the transition between two given states is composed of all the ways in which A-1 nucleons can form a common parent in the initial and final states while the last nucleon changes its state. The probabilities of admixture of these single particle wave functions into the initial and final states decrease as the number of states into which admixture can occur increases. Therefore, the total radiation width may be expected to be proportional to the spacing of levels $\langle D_J \rangle$ near the radiation state multiplied by the sum of all the individual transition probabilities.

Summations of this kind have been carried out by the writer (Cameron 1957*a*, 1957*d*), who found that total radiation widths (in millielectron volts) could be approximately fitted by

$$\Gamma_\gamma = A^{2/3}\langle D_0(U)\rangle \int_0^U E^3 dE / \langle D_0(U-E)\rangle. \tag{12}$$

In evaluating this formula level spacings calculated from equation (9) were used, and the integration was performed numerically for all nuclei whose radiation widths are plotted in Fig. 2.

The ratios of the observed radiation widths to those calculated from equation (12) are plotted in Fig. 3. It may be seen that the shell structure effects evident in Fig. 2 have disappeared in Fig. 3. However, the points in Fig. 3 do appear to define a structure, the trend of which has been drawn in the figure. The following interpretation may be suggested for this structure.

The admixture of single particle wave functions into the actual wave functions at the region of excitation corresponding to the neutron binding energy is largest in the vicinity of the appropriate maxima of the neutron strength functions. According to simple optical models of the nucleus the even partial neutron waves have strength function maxima in the same regions of mass number, and the odd partial waves have strength functions at intermediate regions of mass number. If a nucleus absorbs an *s*-wave neutron, the wave function of the compound state may be thought of as a combination of a large subset of the excited states of the target nucleus which combine with the *s*-wave neutron to preserve the total angular momentum of the state. Certain other excited states of the target nucleus couple *d*-, *g*-, . . . wave neutrons in such a way as also to preserve the total angular momentum and parity of the compound state. Still further representations of the compound state may be obtained by coupling the excited states of that parent nucleus, which has one more neutron and one less proton than the target nucleus, to *s*-, *d*-, *g*-, . . . wave protons. In general the fractional admixtures of these wave functions into that of the compound state probably diminish as the amount of recoupling of the nucleons increases, since extensive recouplings usually represent highly excited states.

The electric dipole transition matrix elements are proportional to the amount of overlap of these parent cores between the initial and final states. The overlap integrals will be larger when the admixtures of single particle wave functions into the initial states are larger, such as will occur in the regions of the *s*-wave neutron strength function peaks ($A \sim 55$ and 160), the *d*-wave neutron peaks ($A \backsim 50$ and 200), and the *s*- and *d*-wave proton peaks ($A \backsim 75$ and 50 respectively). It may be seen in Fig. 3 that the curve is generally in accord with these expectations.

It is also necessary to consider the strengthening of the transition matrix elements owing to increased admixtures of single particle wave functions into the final states. Most of the transitions go to states with roughly half the neutron binding energy (more than half of the neutron binding energy in the case of very heavy nuclei). Since the single nucleon changes its orbital angular momentum in the radiative transition, *s*-wave neutron capture should be enhanced when the middle range of excited states has large admixtures of *p*-, *f*-, *h*-, . . . wave nucleons. Again the lower partial waves should be more important than the higher ones owing to the fact that more extensive recouplings correspond in general to higher excitation energies. This enhancement occurs in between the strength function peaks for the odd partial wave

nucleons, in the same general region as the strength function peaks for the even partial wave nucleons.

These considerations lead to the suggestion that the radiation widths associated with p-wave neutron capture may be somewhat different from those associated with s-wave capture. The broad structure peaks should occur in different regions of mass number. There is as yet insufficient experimental data to test this prediction. The meager information available about radiation widths associated with p-wave capture is also plotted in Fig. 3. Only one of these points is a direct experimental measurement (Saplakoglu, Bollinger, and Cote 1958); the others have been deduced by the writer from the capture cross section curves measured in the kev region at Duke University (Bilpuch and Newson 1958). These latter points are subject to considerable uncertainty.

Some of the increased radiation probability shown in Fig. 3 near $A = 160$ and $A = 240$ may also possibly be due to the fact that nuclei in these regions

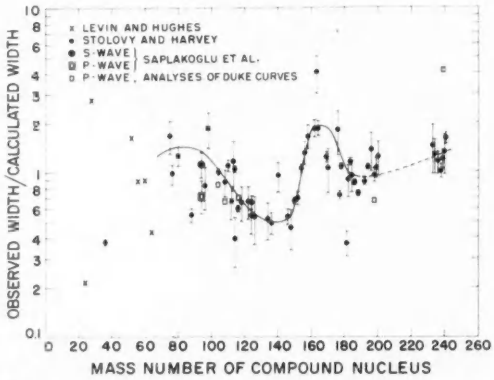


FIG. 3. Ratios of observed radiation widths to those calculated from equation (12). The solid line is intended as a guide to the eye to denote the trend of the points with mass number.

are deformed. It has been found (Fuller, Petree, and Weiss 1958) that the giant dipole resonance is broadened in deformed nuclei. The matrix elements under discussion in this paper might receive some strengthening from collective effects in the wings of the giant resonance.

It therefore appears that equation (12) gives a fairly good account of the total radiation widths. Deviations from this relation are relatively small, are regular in nature, and are probably susceptible to theoretical interpretation. It would be very desirable to check also to see whether the integrand in equation (12) describes the shape of the primary photon spectrum correctly. Unfortunately any experiment has to measure the total photon spectrum which includes a large number of secondary photons. The writer has made cascade calculations designed to convert the primary spectrum of the integrand of equation (12) into the total spectrum which would be measured. These

have been compared with measured spectrum shapes (Bartholomew, Campion, Knowles, and Manning 1957; Cameron and Bartholomew, unpublished). Fairly good agreement has been observed between the calculated and measured spectra, particularly in the heavier nuclei where these statistical methods are more reliable. These comparisons therefore give additional confidence that equation (12) provides a fairly good description of radiation widths.

From this discussion it will be apparent that the results of the Mike test imply that the following relation holds among the heavy uranium nuclei:

$$\langle \Gamma_\gamma \rangle / \langle D_0 \rangle \approx \text{const.} \int_0^U E^3 dE / \langle D_0(U-E) \rangle \approx \text{constant.} \tag{13}$$

EXCITATION ENERGIES

The quantity $\int_0^U E^3 dE / \langle D_0(U-E) \rangle = I$ of equation (13) depends very strongly on the effective excitation energy U. Most empirical and semiempirical mass formulas (see for example Green 1954; Levy 1957; Cameron 1957*b*) predict that $U \approx 2.5$ to 3.0 Mev in U^{254} as compared with 4.7 Mev in U^{238}. The major contribution to the integrand of I occurs at an energy somewhat less than U. If these excitation energies are correct, then the values of I for U^{239} and U^{255} will be about the same only if the level spacing near $U = 2$ Mev in U^{255} is approximately the same as that near $U = 4$ Mev in U^{239}. As may be seen from equation (9) this in turn requires that the density of nucleon orbits near the Fermi level in U^{255} should be approximately double the similar density in U^{239}.

The greater part of the density of nucleon orbits in very heavy nuclei is contributed by the density of neutron orbits. It may be seen from an inspection of the numbers in the appendix of the writer's paper on level spacings (Cameron 1958*a*) that there does not seem to be a significant decrease in neutron orbit spacings for nuclei along the valley of beta stability beyond U^{238}. Hence it is unlikely that any such effect is associated with shell structure.

In the Fermi gas model of the nucleus a doubling of nucleon orbit densities would occur if the nuclear volume were doubled. It is perhaps worth inquiring as to whether any such effect could be associated with the smaller neutron binding energies in U^{255}. We would require a radius expansion of 26%. About 2% is given by the increase in nuclear volume in going from U^{239} to U^{255}.

A reasonable shape for the nuclear potential is (Wilets 1958):

$$V = V_0/(1+\exp(r-R)/a), \tag{14}$$

where $V_0 \approx -40$ Mev, $a \approx 0.65$ fermis (i.e. units of 10^{-13} cm), R is the radius of the half-central-density point, and r is the radial distance corresponding to V. This potential lies outside the nuclear matter, but if the binding energy of the last neutron decreases, then the nuclear matter will extend farther out within this nuclear potential. On this account we may expect the nuclear radius associated with the Fermi level to increase by 5 or 6% in going from U^{239} to U^{255}.

Clearly these increases of nuclear radius are inadequate to provide a doubling of nucleon orbit densities in U^{255} relative to U^{239}. Hence it seems necessary to

conclude that the mass formulas are incorrect, and that the neutron binding energy in U^{255} is nearly as large as that in U^{239}.

In this connection it is perhaps useful to call attention to a possible correction to the mass formula which has been discussed by Szamosi and Ziegler (1956) and by Green (1958). This is the direct symmetry term. Szamosi and Ziegler showed that if the neutron radius R_n and the proton radius R_p are different, then there should be a mass contribution of the form

$$\Delta M = \text{const.}\ [(R_p^2 - R_n^2)/R_p R_n]\ (N-Z), \tag{15}$$

where N and Z are the numbers of neutrons and protons in the nucleus.

Nuclei along the valley of beta stability have approximately equal neutron and proton binding energies, and it appears that for such nuclei $R_p \approx R_n$ (Wilets 1958). Hence the term of equation (15) would make a negligible contribution to nuclei near the valley of beta stability to which the usual mass formulas are fitted. However, in the neutron rich region the argument given above suggests that the effective neutron radius may be significantly larger than the effective proton radius. In this case the term in equation (15) may cause a significant decrease in the masses in the neutron rich region, thus causing neutron binding energies to decrease more slowly than predicted by the usual mass formulas as neutrons are added to a heavy nucleus.

Another possibility which the writer finds more appealing would be the introduction of a Coulomb surface correction term. It has been indicated above that the density distribution of neutrons near the nuclear surface becomes more diffuse as the neutron binding energies decrease. Associated with this additional diffuseness there must be a slight enlargement of the nuclear potential energy distribution. It is probable that this induces an additional, smaller but significant, diffuseness in the surface density distribution of the protons. The Coulomb energy of a heavy uranium nucleus is of the order of 1000 Mev. If the writer's trapezoidal charge model of the nucleus (Cameron 1957*b*) is modified by increasing the surface thickness by only 10^{-14} cm, while keeping the central density constant, the Coulomb energy of a heavy uranium nucleus would decrease by 7 Mev. This would be quite adequate to stabilize the neutron binding energies in the manner required by the conclusions of this paper.

From this discussion it appears that measurements of yields resulting from multiple neutron capture can give significant new useful information about nuclear structure in the unexplored realm of nuclei on the neutron rich side of the valley of beta stability. This information is of great interest in itself, and is needed for a detailed understanding of the process of neutron capture on a fast time scale which is necessary for nucleogenesis and which may occur in type I supernovas (Cameron 1957*c*; Burbidge, Burbidge, Fowler, and Hoyle 1957; Cameron 1958*b*). In this connection it is hoped that more information will be released on the measurements of the yields of the heavy rare earth activities resulting from multiple neutron capture in the fission products made during the Mike explosion, and also on any other measurements which relate to multiple neutron capture in the structural material of the fusion device.

It would also be very useful to conduct further experiments on multiple neutron capture using fusion devices with a deuterium–tritium mixture.

The writer is indebted to G. A. Bartholomew, H. A. Bethe, J. E. Lynn, and A. M. Lane for useful discussions of nuclear radiation widths, to A. E. S. Green for discussions about nuclear masses, and to L. G. Elliott for valuable conversations concerning the interpretation of the Mike test yields.

REFERENCES

BARTHOLOMEW, G. A., CAMPION, P. J., KNOWLES, J. W., and MANNING, G. 1957. Proceedings of the International Conference on the Neutron Interactions with the Nucleus, Columbia University. U.S. Atomic Energy Commission Report TID-7547.

BETHE, H. A. 1958. U.S. Atomic Energy Commission Report LA-2000.

BILPUCH, E. G. and NEWSON, H. W. 1958. To be published.

BLATT, J. M. and WEISSKOPF, V. F. 1952. Theoretical nuclear physics (John Wiley & Sons, Inc., New York).

BURBIDGE, E. M., BURBIDGE, G. R., FOWLER, W. A., and HOYLE, F. 1957. Revs. Modern Phys. **29**, 547.

CAMERON, A. G. W. 1957*a*. Can. J. Phys. **35**, 666.

——— 1957*b*. Can. J. Phys. **35**, 1021.

——— 1957*c*. Chalk River Report CRL-41.

——— 1957*d*. Proceedings of the International Conference on the Neutron Interactions with the Nucleus, Columbia University. U.S. Atomic Energy Commission Report TID-7547.

——— 1958*a*. Can. J. Phys. **36**, 1040.

——— 1958*b*. Ann. Rev. Nuclear Sci. **8** (in press).

FULLER, E. G., PETREE, B., and WEISS, M. S. 1958. To be published.

GHIORSO, A., THOMPSON, S. G., HIGGINS, G. H., SEABORG, G. T., STUDIER, M. H., FIELDS, P. R., FRIED, S. M., DIAMOND, H., MECH, J. F., PYLE, G. L., HUIZENGA, J. R., HIRSCH, A., MANNING, W. M., BROWNE, C. I., SMITH, H. L., and SPENCE, R. W. 1955. Phys. Rev. **99**, 1048.

GREEN, A. E. S. 1954. Phys. Rev. **95**, 1006.

——— 1958. Revs. Modern Phys. **30**, 569.

GRYZINSKI, M. 1958. Phys. Rev. **111**, 900.

HUGHES, D. J. and SCHWARTZ, R. B. 1957. Supplement No. 1. U.S. Atomic Energy Commission Report BNL-325.

HUIZENGA, J. R. and DIAMOND, H. 1957. Phys. Rev. **107**, 1087.

LANE, A. M. and LYNN, J. E. 1957. Proc. Phys. Soc. (London), A, **70**, 557.

LEVIN, J. S. and HUGHES, D. J. 1956. Phys. Rev. **101**, 1328.

LEVY, H. B. 1957. Phys. Rev. **106**, 1265.

MOTTLESON, B. R. and NILSSON, S. G. 1958. To be published.

NEWTON, T. D. 1956. Can. J. Phys. **34**, 804.

POST, R. F. 1956. Revs. Modern Phys. **28**, 338.

SAPLAKOGLU, A., BOLLINGER, L. M., and COTE, R. E. 1958. Phys. Rev. **109**, 1258.

SEABORG, G. T. 1958. Evening Lecture, Second United Nations Conference on the Peaceful Uses of Atomic Energy, Geneva, September 9.

STOLOVY, A. and HARVEY, J. A. 1957. Phys. Rev. **108**, 353.

SZAMOSI, G. and ZIEGLER, M. A. 1956. Acta Phys. Acad. Sci. Hung. **6**, 345.

VASILYEV, YU. A., ZAMYARTNIN, YU. S., TOROPOV, P. V., and FOMUSHKIN, E. F. 1957. Atomnaya Energ. **3**, 542.

WILETS, L. 1958. Revs. Modern Phys. **30**, 542.

THE TRANSPORT OF HEAT BETWEEN DISSIMILAR SOLIDS AT LOW TEMPERATURES[1]

W. A. LITTLE[2]

ABSTRACT

The resistance offered to the flow of heat by the mismatch of the elastic constants at the interface between two materials has been calculated. It is shown that for a perfectly joined interface the heat flow is proportional to the difference of the fourth powers of the temperature on each side of the interface. Deviations from this temperature dependence are to be expected for rough surfaces and for surfaces pressed into contact with one another. The calculated contact resistance between some common solids is given, and graphs are presented from which the heat flow between any two materials may be computed. It is shown too that the spin–phonon, phonon–electron, and phonon–phonon relaxation processes give rise to additional resistive processes in some solids, some of which restrict the heat flow at the surface. The theoretical results compare well with the available experimental data. However, the problem of the contact resistance between helium and metals is still unresolved. A tentative explanation of this is presented which predicts that there should be an appreciable difference between the contact resistance of a metal in the normal and in the superconducting state.

INTRODUCTION

In experiments conducted at very low temperatures it has become well established that a finite difference of temperature will occur at an interface between two dissimilar materials when a heat flux is maintained across this interface. Kapitza (1941) first observed such an effect between helium II and a metal, and since then many investigators have reported similar results with helium II (Fairbank and Wilks 1955) and helium 3 (Fairbank and Lee 1957) and between paramagnetic salts and metallic fins (Goodman 1953). Apart from several attempts to explain the contact resistance in the specific case of a helium–metal interface (Kronig and Thellung 1950; Gorter, Taconis, and Beenakker 1951; and Khalatnikov 1952), no general explanation of the effect has been offered. In view of the fact that this resistance plays a very important role in all experiments using cooling by contact at temperatures below 1° K, a general explanation is long overdue. In this paper such an attempt has been made, firstly, by considering the ideal case of two perfectly joined media and, secondly, by extending this result to several cases of practical importance.

I. PERFECT JUNCTION

Let us consider two semi-infinite isotropic solids joined in the plane zy at $x = 0$. In order to simplify the problem let us suppose that the modulus of rigidity of both materials is zero so that we neglect the transverse phonons in both solids. This greatly simplified the boundary conditions at the interface

[1]Manuscript received November 14, 1958.
Contribution from the University of British Columbia, Vancouver, British Columbia.
[2]Present address: Stanford University, Stanford, California, U.S.A.

and yet allows one to see the trend of the solution in the more general case considered in II.

The number of phonons of energy $h\nu$ incident upon a small area dA of the interface per unit time at an angle of incidence between θ_1 and $\theta_1+d\theta_1$ will be:

$$\tfrac{1}{2}c_1N_1(\nu)\cos\theta_1\sin\theta_1 d\theta_1 dA d\nu, \tag{1}$$

where c_1 is the group velocity of the longitudinal phonons in the first medium and $N_1(\nu)d\nu$ is the number of phonons of energy lying between $h\nu$ and $h(\nu+d\nu)$ per unit volume of the first medium.

Some of these phonons will be reflected at the interface and some will be refracted into the second medium at an angle of refraction between θ_2 and $\theta_2+d\theta_2$. The two angles θ_1 and θ_2 are related by Snell's law of acoustic refraction from which one may obtain the relation:

$$d\theta_2 = \frac{c_2\cos\theta_1}{c_1\cos\theta_2}\cdot d\theta, \tag{2}$$

where c_2 is the group velocity of longitudinal phonons in the second medium. The nett flow of heat across the interface from medium 1 to medium 2 due to these phonons leaving or entering at an angle θ_1 is then:

$$\frac{dQ}{dt} = \tfrac{1}{2}\iiint N_1(\nu)h\nu c_1\alpha_1(\theta_1)\cos\theta_1\sin\theta_1 d\theta_1 d\nu dA - \tfrac{1}{2}\iiint N_2(\nu)h\nu c_2\alpha_2(\theta_2)\cos\theta_2\sin\theta_2 d\theta_2 d\nu dA. \tag{3}$$

In this expression $\alpha_1(\theta_1)$ is the transmission coefficient of the interface for phonons incident at an angle θ_1 in medium 1. Because the longitudinal phonons are the quantized longitudinal acoustic waves, this coefficient will be given by the equivalent expression for acoustic waves (Rayleigh 1894):

$$\alpha_1(\theta_1) = \alpha_2(\theta_2) = \frac{\dfrac{4\rho_2c_2}{\rho_1c_1}\cdot\dfrac{\cos\theta_2}{\cos\theta_1}}{\left(\dfrac{\rho_2c_2}{\rho_1c_1}+\dfrac{\cos\theta_2}{\cos\theta_1}\right)^2}, \tag{4}$$

where ρ_1 and ρ_2 are the densities of the two media.

The value of $N(\nu)$ is given by the number of degrees of freedom of the lattice lying between ν and $\nu+d\nu$, and the Bose–Einstein energy distribution.

$$N(\nu)d\nu = \frac{4\pi}{c_1^3}\frac{\nu^2 d\nu}{(e^{h\nu/kT}-1)}. \tag{5}$$

Inserting (5) in (3) and integrating with respect to A one obtains the expression:

$$\frac{dQ}{dt} = \frac{2\pi k^4\Gamma A}{h^3c_1^2}\left[T_1^4f(T_1)-T_2^4f(T_2)\right]\text{ ergs/second} \tag{6}$$

where

$$f(T) = \int_0^{h\nu_m/kT} \frac{z^3 dz}{e^z - 1} \tag{7}$$

with

$$z = h\nu/kT$$

and

$$\Gamma = \int_0^{\pi/2} \alpha_1(\theta_1) \sin\theta_1 \cos\theta_1 \, d\theta_1.$$

At low temperatures one may make the usual approximation ($h\nu_m/kT = \infty$) in the integral over z as it occurs in the Debye theory of specific heats, and thus obtain the relation:

$$\frac{dQ}{dt} = \frac{2\pi k^4 \Gamma A}{h^3 c_1^2} \left(\frac{\pi^4}{15}\right) (T_1^4 - T_2^4) \tag{8}$$

$$= 5.01 \times 10^{16} \frac{\Gamma A}{c_1^2} (T_1^4 - T_2^4) \text{ ergs/second}$$

with c_1 in cm/sec and T in ° K.

For small values of the ratio $(T_1 - T_2)/T_1$ we may write this in the convenient form:

$$\frac{dQ}{dt} = 2.0 \times 10^{17} \frac{\Gamma A}{c_1^2} T_1^3 \, \Delta T \text{ ergs/second} \tag{9}$$

where

$$\Delta T = T_1 - T_2.$$

The integral Γ is a complicated function of the densities and of the acoustic velocities of the two media. In its explicit form it is difficult to use, so it has been computed for a range of values of the ratio of the densities and the ratio of the acoustic velocities. The integral is given in graphical form in Fig. 1.

This simple acoustic model shows that one should expect a discontinuous change of temperature across the interface and that for small heat fluxes the drop in temperature ΔT should be proportional to the heat flux and to the third power of the mean absolute temperature. From the variation of Γ as shown in Fig. 1, with respect to the relative densities and acoustic velocities in the two media, it can be seen that to obtain the maximum transfer of heat from one medium to another the densities should be similar and the acoustic velocities similar. The strong dependence of Γ on the relative acoustic velocities for $c_2 - c_1$ is due to the total reflection of phonons for θ_1 greater than the critical angle $\sin^{-1}(c_1/c_2)$. It should be noted that the matching of the specific acoustic impedance of each medium (i.e., $\rho_1 c_1 = \rho_2 c_2$) is not the condition for the maximum transfer of heat when $c_2 > c_1$.

The maximum value of Γ is 0.5, which corresponds to a perfect match of the densities and the acoustic velocities. In this case (6) predicts that a

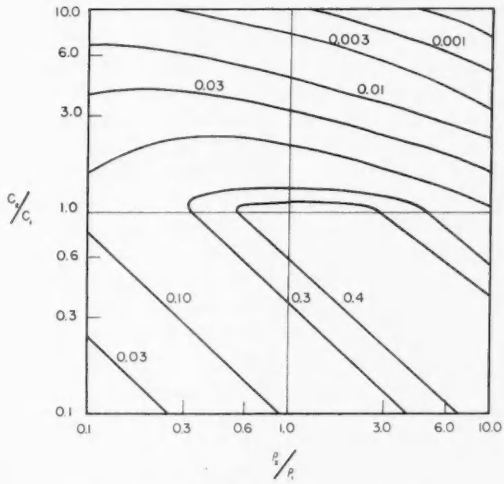


FIG. 1. Contours of Γ for transmission of heat from medium 1 to medium 2 with respect to the ratio of the acoustic velocities c_2/c_1 and the densities ρ_2/ρ_1 of the two media. Poisson's ratio is equal to 0.5 in both media.

temperature difference will occur across the boundary. This apparent paradox is explained by the fact that it was tacitly assumed in (6) that the two media were each at a uniform temperature in the presence of the heat flux. This implies that the region in which the temperature is measured should extend for a distance no greater than the mean free path of the phonons on each side of the barrier. In the particular case of identical media sustaining a heat flux the theory of thermal conductivities requires that there should be a temperature drop over a distance of the mean free path.

That this is so can be shown as follows. If $c_1 = c_2$ and $\rho_1 = \rho_2$ then $\alpha_1(\theta) = 1$ for all values of θ. Let T be the temperature at the barrier and let the mean free path of phonons in this medium be λ then the temperature T_1 of a point where a phonon is last scattered before striking the barrier will be $[T+(\partial T/\partial z)\lambda\cos\theta]$ and likewise T_2 will be $[T-(\partial T/\partial z)\lambda\cos\theta]$. At low temperatures and for small values of T_1-T_2 we obtain

$$\frac{dQ}{dt} = \frac{2\pi k^4 A}{h^3 c_1^2}\left(\frac{\pi^4}{15}\right) 4T_1^3 \frac{\partial T}{\partial z} 2\lambda \int_0^{\pi/2} \cos^2\theta \sin\theta \, d\theta$$

$$= \frac{1}{3} C_v \lambda c_1 \left(\frac{\partial T}{\partial z}\right) A,$$

where C_v = lattice specific heat unit/volume which gives a thermal conductivity of $\frac{1}{3}C_v\lambda c_1$, which is the usual expression for the conductivity at low temperatures.

The conduction of heat over these small distances at very low temperature is, of course, analogous to that in gases at low pressures. In the case of a gas

the heat flow for distance less than the mean free path is limited by the number of molecules, that is, the pressure, while in a solid the heat flow is limited by the number of phonons, which depends upon the temperature. Hence the contact resistance is related to the fact that for distances less than the mean free path the heat flux is not proportional to the temperature gradient but requires a finite temperature difference. The effect of the interface is to reduce this heat flux for a given temperature difference in proportion to the transmission coefficient of the barrier.

It is possible to extend this acoustic model from the ideal case considered above to obtain a qualitative understanding of the effect of roughness and of imperfect contact at the interface upon the temperature dependence and magnitude of the contact resistance. This may be seen by first considering what happens if the effective area of the interface depends upon the frequency of the incident lattice waves. Suppose that the effective area of the interface $A'(\nu)$ for phonons of frequency ν may be written as

$$A'(\nu) = A_0\nu^n, \tag{10}$$

where n may be 0 or a positive or negative quantity. We will assume that such an expression is possible over the range of frequencies of phonons which contribute appreciably to the heat content of the lattice. Using (10), equation (6) now yields

$$\frac{dQ}{dt} = B\,(T_1^{n+4} - T_2^{n+4}) \tag{11}$$

where B contains terms independent of the temperature.

The heat flux across the interface now will be more or less strongly dependent upon the absolute temperature depending on n being positive or negative.

Let us now consider two cases of practical interest, firstly, an interface which is rough but perfectly joined at all points and, secondly, an interface at which contact between the two media occurs at some points but not at others.

The Rough Interface

Let Fig. 2 represent the interface between the two rough media. Let $\bar{A}$ be defined as the mean amplitude of the roughness. If the mean free path of the

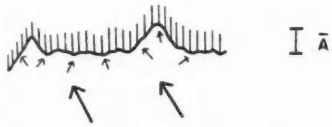


FIG. 2. Diagrammatic representation of a rough interface illustrating the larger effective area of the interface to short lattice waves compared with that for long waves.

phonons is comparable to $\bar{A}$ then the effective area of the interface will differ for phonons of different frequency. The roughness of the interface will occur over a small fraction of a wavelength of a long lattice wave and, consequently,

the wave will be transmitted as if the interface was smooth. For these waves the effective area will be the macroscopic area of a plane at the interface. On the other hand the short lattice waves will be transmitted across each individual hill and dale at the interface and the effective area will be the integrated microscopic area of the interface. We should expect, then, the effective area to increase with increasing phonon energy.

$A'(\nu)$ now will be given by (10) with a positive value of n. Also, the transport of heat across the interface will be greater than that given by equation (8) due to the average effective area being greater than the macroscopic area of the interface.

If, however, the mean free path is much greater than $\bar{A}$ the number of phonons of given energy incident upon the interface will be determined by the cross-sectional area of the solid normal to the heat flux and not by the microscopic area of the interface. In this case the effective area of the interface will be the same for all phonons and equation (8) will apply with A given by the macroscopic cross-sectional area of the interface.

Imperfect Contact

Another case of some practical interest is that of the imperfect junction of two media. In many experiments of contact cooling at temperatures below 1° K, the specimen is cooled via a metallic link which is pressed into contact with a powdered paramagnetic salt. Berman (1956) has shown that the amount of heat transported across the interface between two surfaces pressed into contact is proportional to the total force applied. This suggests that the contact occurs at a few points, and that as greater pressure is applied the surfaces deform and more points make contact, the number of point contacts then being proportional to the applied pressure with each carrying approximately the same heat flux. In the case of a paramagnetic salt being pressed in powder form against a metallic fin the contact would occur at a number of points. It is instructive to apply our acoustic model to this case.

We will suppose that the interface can be represented by Figs. 3(*a*) and 3(*b*)—a series of point contacts separated by a mean distance, $\bar{B}$. Phonons with wavelengths which are very large compared with $\bar{B}$ will move all the points lying adjacent to one another approximately in phase, and as a result the whole boundary between adjacent points will move in phase (see Fig. 3(*a*)). The mean displacement at points on the boundary not in contact with the first medium will be only slightly less than that at points of contact. We should expect, therefore, that the strain in the second medium a short distance from the interface will be only slightly less than it would be if the junction were perfect. Hence the energy transported across the interface, which is proportional to the square of the strain, will be only slightly less than that for the ideal case. On the other hand, for phonons with wavelengths which are small compared with $\bar{B}$ the motion of adjacent points of contact will differ by a random phase factor. This factor will be determined for each point by the spatial localization of the phonon and the actual value of the separation of the points of contact. In this case the motion of the barrier *not* in contact

with the first medium will tend to zero due to the destructive interference of the surface waves from all the nearby points of contact (see Fig. 3(*b*)). The mean displacement in the second medium now will be much smaller and the transmitted energy correspondingly reduced.

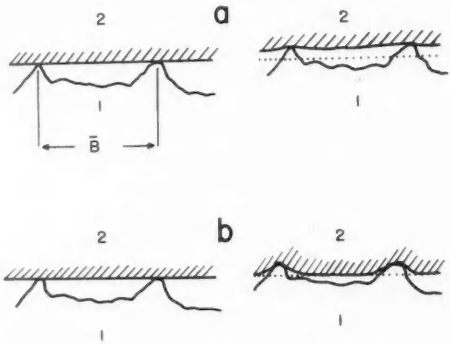


FIG. 3. Diagrammatic representation of the interface of two solids in imperfect contact.
(*a*) The surface is shown, firstly, at rest, and, secondly, at the peak of a low frequency lattice wave. The wavelength in this case is much greater than the mean separation of the points of contact $\bar{B}$.
(*b*) The corresponding picture for a high frequency lattice wave. The dotted line represents the surface of the second medium at rest.

From the above considerations one would expect the effective area of the interface to decrease with increasing phonon frequency, giving (10) with a negative n. Consequently, the transport of heat across the barrier would be less temperature-dependent and the total heat flux less than that given by (8) for the macroscopic area of contact.

II. THE GENERAL CASE

In Section I we have considered the simple case of the transport of heat between two media in which the moduli of rigidity are both zero. This is a somewhat artificial case so it is necessary, for practical reasons, to consider the more general case where both longitudinal and transverse phonons occur. We shall consider the case of a perfect junction. The reasoning for the rough and imperfect junction is equally applicable to this case so there is no need to reconsider these effects.

Equation (8) may be readily generalized to add the contribution from the transverse phonons to the heat flow across the interface. We then obtain:

$$\frac{dQ}{dt} = 5.01\times10^{16}A\left[\frac{\Gamma_l}{c_l^2}+\frac{2\Gamma_t}{c_t^2}\right](T_1^4-T_2^4)\ \text{ergs/sec cm}^2, \tag{12}$$

where Γ_l = transmission coefficient of longitudinal phonons,
Γ_t = transmission coefficient of transverse phonons,
c_l = velocity of longitudinal phonons,
c_t = velocity of transverse phonons.

In this case the detailed evaluation of Γ_l and Γ_t is much more difficult. The reason for this is that at the interface each incident wave, whether longitudinal or transverse, in general breaks up into four waves, reflected transverse and longitudinal waves, and refracted transverse and longitudinal waves (Kolsky 1953). Because of the different velocities of the transverse and longitudinal waves, two critical angles of reflection are found and in the interval between these two angles a phase shift occurs in the refracted and reflected waves whose magnitude depends upon the angle of incidence. In this case there is no simple expression corresponding to equation (4). It is easier to calculate the transmission coefficient for each angle of incidence by a numerical solution of the equations determined by the boundary conditions. This has been done and the values of Γ_l and Γ_t have been computed for a range of relative densities, Poisson's ratio of the media, and the relative velocities of the longitudinal waves.

Method of Calculation

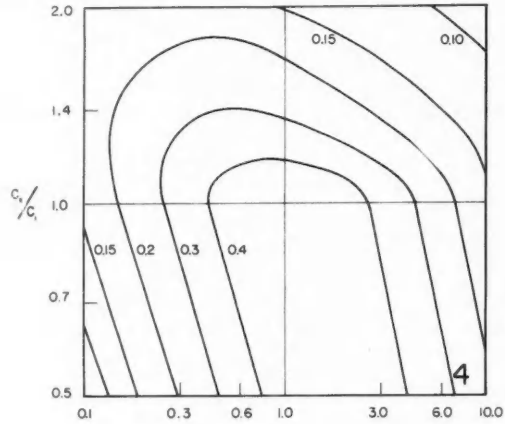
The ALWAC IIIE computer at the U.B.C. Computing Center was programed to solve the eight linear equations given by the boundary conditions for the amplitude and phase of each of the four possible reflected and refracted waves (Kolsky 1953). The fraction of the energy of the incident wave transmitted to the second medium was computed for each angle of incidence θ. The integral was then performed using a Legendre–Gauss five point quadratures routine which approximates the transmission coefficient to a polynomial of ninth degree in θ. The program was checked at several points using the reciprocity relation:

$$\left[\frac{\Gamma_l}{c_l^2}+\frac{2\Gamma_t}{c_t^2}\right]_{ij} = \left[\frac{\Gamma_l}{c_l^2}+\frac{2\Gamma_t}{c_t^2}\right]_{ji} \tag{13}$$

where the subscripts indicate the values of Γ_l and Γ_t within the bracket from medium i to medium j.

The contours of Γ are given for a range of relative densities of the two media and the ratios of the velocities of longitudinal waves in the two media. In Figs. 4 and 5 Poisson's ratio σ was chosen to be 0.33 in both media, corresponding to $c_l/c_t = 2$; while in Fig. (6) and (7), $\sigma = 0.33$ for the first medium and 0.44 for the second. Poisson's ratio = 0.44 corresponds to $c_l/c_t = 3$. The values of Γ_t given in Figs. 5 and 7 are the values obtained after integrating over all possible angles of polarization of the transverse waves.

In Table I Γ_t, Γ_l, and dQ/dt are given for some common solids. These have been computed from the room temperature densities and acoustic velocities as given in the *American Institute of Physics Handbook*. In the case of crystalline solids these acoustic velocities are the values averaged over the different crystal directions. It has been assumed that we can approximate these solids to isotropic solids with these average values of the room temperature acoustic velocities and densities. In many cases this is not a very good approximation but the values of Γ still will be correct to within 10 to 20% judging from the variations of Γ in Figs. 4, 5, 6, and 7. This accuracy is sufficient to allow



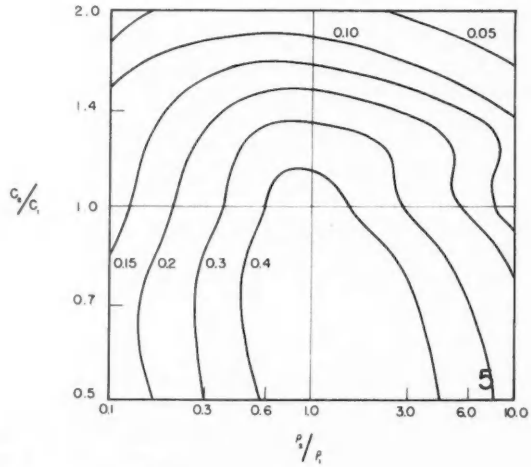
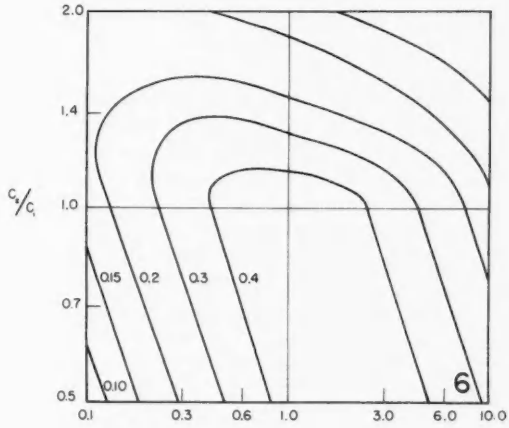


FIG. 4. Contours of Γ_l for the transmission of heat from medium 1 to 2 with respect to the ratio of the velocity of longitudinal waves in the two media and the relative densities. Poisson's ratio is equal to 0.33 in both media.

FIG. 5. Contours of Γ_t for the transmission of heat from medium 1 to 2 with respect to the ratio of the velocities of longitudinal waves in the two media. Poisson's ratio is equal to 0.33 in both media.

these values to be used as a guide in selecting suitable materials for optimum heat transfer.

So far we have only considered the contact resistance due to the imperfect transmission of phonons across the interface. In certain specific cases such as the transport of heat from a metal to a dielectric solid or from a metal to a paramagnetic salt there are additional factors which can contribute to the



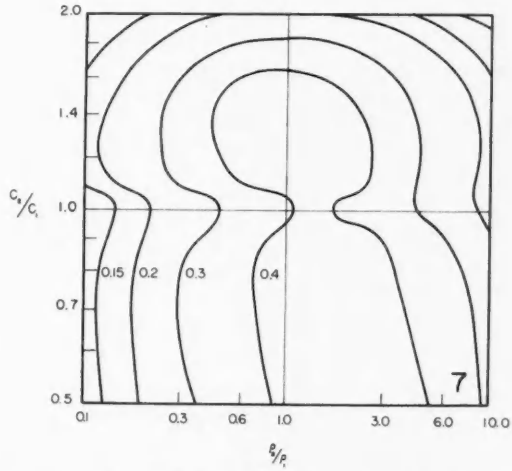


FIG. 6. Contours of Γ_l for Poisson's ratio equal to 0.33 in the first medium and 0.44 in the second.

FIG. 7. Contours of Γ_t for Poisson's ratio equal to 0.33 in the first medium and 0.44 in the second.

apparent contact resistance. This resistance does not necessarily occur at the points of contact of the two media but experimentally one might have difficulty distinguishing these effects from the true contact resistance discussed above. In the case of a metal in contact with a dielectric solid one has to consider the resistive contribution due to the finite electron–phonon relaxation time, whereas in the case of a paramagnetic salt in contact with a metal one has in addition to the phonon–electron relaxation time the spin–lattice and phonon–phonon relaxation times. The contribution of these will be discussed in turn.

III. PHONON–ELECTRON CONTACT RESISTANCE

Let us calculate the rate dQ/dt at which heat may be transported from the system of phonons in a metal to the conduction electrons. Let $L(\nu)$ be the mean free path of a phonon of energy $h\nu$ for collisions with the conduction electrons. We will suppose that we are concerned with very low temperatures so that the mean free path for phonon–phonon collisions is very great compared with $L(\nu)$. In the simple theory of electron–phonon interactions there is no coupling between the transverse phonons and the conduction electrons so we shall calculate the transport of heat from the longitudinal phonons to the conduction electrons alone. This will serve to give us the order of magnitude of this effect. Furthermore, we will assume that on the average a fraction α of the energy of a phonon can be transferred to the electrons at a collision. We obtain the expression:

$$\frac{dQ}{dt} = \int_{\nu} \frac{\alpha h\nu c_1 N(\nu) d\nu}{L(\nu)}. \tag{14}$$

The mean free path in a metal may be calculated in terms of the electron–phonon interaction and is given by Wilson (1953) as:

$$\frac{1}{L(\nu)} = \left(\frac{h\nu}{k\Theta}\right) \frac{m^{*2} a C_j^2}{h^2 k\Theta M} \left(\frac{4\pi}{3}\right)^{1/3} 6\pi^2 \tag{15}$$

where m^* = effective mass of conduction electrons,
a = lattice constant,
C_j = electron–phonon interaction energy 1 ev $< C_j <$ 10 ev,
Θ = Debye temperature,
M = mass of an atom of the lattice.

Putting (15) in (14) and using the Bose–Einstein distribution for $N(\nu)$ one obtains the nett transfer of heat from the phonons at temperature T_1 to the electrons at temperature T_2

$$\frac{dQ}{dt} = 9R\alpha \left[\frac{m^{*2} a^2 C_j^2}{h^2 M \Theta^4}\right] (T_1^5 - T_2^5)\, 6\pi^2 \left(\frac{4\pi}{3}\right)^{2/3} \times \int_0^\infty \frac{x^4}{e^x - 1}\, dx \text{ ergs/sec g-atom.} \tag{16}$$

This expression has been evaluated for copper on the assumption that $m^* = m$, and $C_j = 3$ ev, and $\alpha = 0.25$

$$\frac{dQ}{dt} = 1.0 \times 10^{10}\, (T_1^5 - T_2^5) \text{ ergs/sec/cc} \tag{17}$$

and the mean free path for the effective transfer of energy to the conduction electrons for phonons of energy $h\nu = kT$

$$L = 4 \times 10^{-3} (1/T) \text{ cm.} \tag{18}$$

These calculations serve to illustrate two points. Firstly, it is necessary to have an appreciable volume as well as a large surface area of a metal in contact

with a cooling salt in order to effect the maximum transfer of heat at very low temperatures. Secondly, the technique of embedding many fine copper wires in a paramagnetic salt to obtain good thermal contact has only limited value at extremely low temperatures. There is no point in using extremely fine wire in order to increase the surface area alone if the diameter of the wire becomes comparable to the mean free path of the phonons. If the diameter is of this order those phonons which do penetrate the surface will have a good chance of escaping without transferring their energy to the conduction electrons.

IV. CONTACT RESISTANCE WITH PARAMAGNETIC SALTS

Where a paramagnetic salt is used to cool a metal in addition to the contact resistance at the surface and that due to the finite phonon–electron relaxation time, there are other resistive elements due to the spin–lattice and the phonon–phonon relaxation times. It appears to be generally accepted now that in most paramagnetic salts the spin–phonon relaxation times are much shorter than the phonon–phonon relaxation times (Strandberg 1958), with the result that the usual equilibrium distribution of the phonons is distorted in the region of phonon energies close to the spin energy.

This has a very important effect upon the transport of heat between the spin system and a metal. This can be shown quite simply by making the approximation that those phonons lying in the frequency range ν_1 to ν_2 at the electron spin frequency are so strongly coupled to the spin system that their distribution between these frequency limits corresponds to that of the spin temperature, while those phonons outside this range are completely decoupled from the spin system and have a distribution corresponding to the lattice temperature. If this salt is in contact with a metal at a different temperature to the spin temperature those phonons outside the range ν_1 to ν_2 will rapidly come to the temperature of the metal because of the small heat capacity of the lattice at these temperatures, while those phonons inside the range ν_1 to ν_2 will remain at the spin temperature and will be the only means by which energy can be transferred between the spin system and the metal. As a consequence of this the thermal conductivity will be lower than that obtained by assuming an equilibrium phonon distribution and the contact resistance higher. In this case we must use the limits $h\nu_1/kT$ and $h\nu_2/kT$ for the integral occurring in (7). We then obtain for the heat flux at the salt–metal interface:

$$\frac{dQ}{dt} = 5.01\times10^{16}\left(\frac{15}{\pi^4}\right)\left[\frac{\Gamma_l}{c_l^2}+\frac{2\Gamma_t}{c_t^2}\right]A\{T_1^4 f(T_1)-T_2^4 f(T_2)\}\ \text{ergs/second} \tag{19}$$

where

$$f(T) = \int_{h\nu_1/kT}^{h\nu_2/kT} \frac{z^3}{(e^z-1)}\,dz. \tag{20}$$

There are several cases of interest in which one can approximate to the value of the integral.

(a) *High Temperature Approximation*

If the temperature is high enough for both $h\nu_1/kT$ and $h\nu_2/kT$ to be much

less than one and low enough to satisfy the conditions for a non-equilibrium phonon distribution, the exponent may be expanded and the integral becomes $\int z^2\,dz$.

Hence,

$$(21)\qquad \frac{dQ}{dt} = 5.01\times10^{16}\left(\frac{15}{\pi^4}\right)\left[\frac{\Gamma_l}{c_l^2}+\frac{2\Gamma_t}{c_t^2}\right]\frac{1}{3}\left(\frac{h}{k}\right)^3(\nu_1^3-\nu_2^3)(T_1-T_2).$$

The heat flux is now much less than that given by (8) and dQ/dT is linearly dependent upon the temperature difference across the interface.

(*b*) *Low Temperature Approximation*

If $\nu_1 = 0$ and $h\nu_2/kT$ is $\gg 1$ we may use the limits 0 and ∞ for z in the integral and obtain the same relation as (8) at very low temperatures. In the region lying between the low and high temperature approximation one should expect the temperature dependence to vary from the fourth power at low temperatures as given by (8) to the linear dependence of (21) at higher temperatures.

(*c*) *High Field Approximation*

If a large magnetic field is applied to the salt, ν_1 and ν_2 will lie close to the electron Larmor precession frequency $\bar{\nu}$ and we will presume that this width $|\nu_1-\nu_2| = \Delta\nu$ will be much less than $\bar{\nu}$. e^z will be large compared to 1 in the denominator of the integral which may then be approximated to give:

$$(22)\qquad \frac{dQ}{dt} = 5.01\times10^{16}\left(\frac{15}{\pi^4}\right)\left[\frac{\Gamma_l}{c_l^2}+\frac{2\Gamma_t}{c_t^2}\right]\left(\frac{h\bar{\nu}}{k}\right)^4\left(\frac{\Delta\nu}{\bar{\nu}}\right)\{e^{-h\bar{\nu}/kT_2}-e^{-h\bar{\nu}/kT_1}\}.$$

In this case we should expect a strongly field and temperature-dependent contact resistance.

It should be noted in the foregoing that the thermal conductivity for the paramagnetic salt will also be affected in the same way for each of the above cases. In particular, the thermal conductivity due to those phonons which interact with the spins will be much less than that due to the bulk of the phonons. Because, firstly, there are fewer phonons able to carry the heat from the spins and, secondly, those that are able are severely scattered by the spins themselves and consequently the diffusion of the spin energy to the surface is much slower than one would expect from the gross thermal conductivity.

At very low temperatures one other limitation to the transfer of heat to or from the spin system might be the true spin–lattice relaxation time. One may optimize this heat transfer by the use of ions with strong spin–orbit coupling such as Cr^{++}, Ti^{++}, and Cu^{++} and avoid those ions in an *S*-state, such as Fe^{+++} and Mn^{+++}.

V. CONTRIBUTION OF SURFACE WAVES

In general one should not expect the surface or Rayleigh waves to play an important role in the transfer of heat from one solid to another. In the first place the total heat flux round the edges of the solid due to these surface waves

is much smaller than through the body of the solid. This is because the waves are localized within about a wavelength of the surface and all the heat must be transmitted through this region. The phonons or body waves are not limited in this way and, consequently, they can carry a far greater heat flux. In the second place the surface waves may be excited by the phonons themselves. However, the two additional interactions which are necessary to transfer the heat energy to and from the surface waves must introduce a further resistive contribution in addition to the boundary conditions which must be satisfied by the surface waves themselves. There are, however, a few exceptional cases where the surface waves may play a more important role. These are where: (i) the surface waves may interact directly with the conduction electrons or with the surface energy states in a metal; or (ii) where the acoustic velocities in the two media are so different that almost all the incident phonons are totally reflected. The totally reflected waves give rise to a surface disturbance in the second medium which can interact directly with the phonons or with the conduction electrons. Because of the small coefficient of transmission of the phonons and the large density of surface waves in this case, the role of each could be expected to be reversed and the latter could become then the dominant transfer mechanism.

Further investigation is required to determine when these waves do, in fact, become the dominant transfer mechanism.

DISCUSSION

From the foregoing analysis one can draw the following conclusions regarding the contact resistance.

(*a*) In general the transfer of heat between two dissimilar materials at low temperatures should not be a more serious problem than the thermal conductivity because the values of Γ seldom are much smaller than 0.2 for common solids. The surface resistance in this case would be equivalent to a little more than the thermal resistance of a thickness of the first medium equal to two phonon mean free paths.

(*b*) Exceptions to these will occur in the following cases where the contact resistance will be abnormally high:

(i) for materials in imperfect contact where the effective area of the bodies in contact is small;

(ii) between media in which there is a large difference in the densities or acoustic velocities where the Γ are small (see Figs. 4, 5, 6, and 7);

(iii) between the spin system of a paramagnetic salt and a solid in contact with it where only a small fraction of the available phonons are able to carry the energy from the spin system. This should be particularly pronounced in the presence of a strong magnetic field.

These conclusions can be compared with the available experimental results. The contact resistance from liquid helium to copper has been measured by several workers. Ambler (1953) obtained the expression $dQ/dt = 3.4 \times 10^5(T_1^4 - T_2^4)$ ergs/sec cm^2. This is the expected temperature dependence but numerically far greater than the calculated value given in Table I. Clearly

TABLE I

From:	To:	Γ_l	Γ_t	dQ/dT in:	ergs/sec cm²
KCr alum	Copper	0.14	0.19	4.7×10^6	$(T_1^4-T_2^4)$
KCr alum	Silver	0.26	0.25	7.3 "	"
KCr alum	Gold	0.17	0.21	5.8 "	"
Pyrex	Copper	0.34	0.36	3.7 "	"
Lucite	Copper	0.11	0.02	2.3 "	"
Garnet	Copper	0.46	0.45	2.3 "	"
Copper	Helium	0.001	0.0008	1.84×10^3	"

the transmission of phonons across the interface alone is quite insufficient to account for the experimental results. However, in this case one should expect the surface waves to be of importance because the ratio of the acoustic velocities is greater than 10. In this particular case we have calculated the amplitude of the surface disturbance caused by totally reflected phonons by solving the boundary conditions as described earlier. It was found that the experimental results then can be explained if about 10% of the energy in this disturbance is transferred to the metal. Such a strong interaction suggests that the surface disturbance interacts directly with the conduction electrons. If this view is correct, there should be an appreciable difference in the contact resistance of a metal in the normal and in the superconducting state. Furthermore, if the surface of the metal is covered with an oxide layer or a layer of dirt only the longer wavelength phonons would be able to penetrate to the conduction electrons. This would give rise to a heat flux proportional to the difference of the third powers of the temperatures of the metal and the helium which would explain the results obtained by Fairbank and Wilks (1955).

It is well known that the transfer of heat between bodies merely pressed into contact with one another is extremely small at low temperatures. Moreover, Berman (1956) found the temperature dependence of this heat flux to be proportional to T^2 at helium temperatures. His results are in agreement with the temperature dependence and magnitude predicted for a pair of surfaces in poor contact treated in Section I.

The results of Section IV provide an explanation of the measurements of Goodman (1953) and more recently, of Miedema, Postma, and Steenland (1958) of the contact resistance from copper to a paramagnetic salt. Goodman measured this contact resistance and found dQ/dT proportional to the square of the absolute temperature. The more recent measurements of Miedema *et al.* on a single crystal of potassium chrome alum cemented to a copper strip show that dQ/dT is proportional to $(T_1^2-T_2^2)$ at high temperatures and $(T_1^4-T_2^4)$ at low temperatures. This agrees with the predicted temperature dependence obtained on the assumption of a non-equilibrium phonon spectrum. At low temperatures the numerical value they obtain was $dQ/dt = 1.2\times10^5(T_1^4-T_2^4)$ergs/sec cm². The calculated value in Table I is about three times greater than this. This is a very satisfactory agreement because the crystal will probably be strained close to the copper surface due to differential contraction and this will impede further the flow of heat.

ACKNOWLEDGMENTS

I wish to thank Dr. J. M. Daniels for helpful advice and criticism during the preparation of this manuscript and Dean G. M. Shrum for the facilities of the Department of Physics during my stay. I am indebted to the staff of the U.B.C. Computing Center for much advice and help in programing the ALWAC IIIE Computer. This work was carried out during the tenure of a National Research Council Postdoctoral Fellowship for which I wish to express my thanks.

REFERENCES

AMBLER, E. 1953. Thesis, University of Oxford.
BERMAN, R. 1956. J. Appl. Phys. **27**, 318.
FAIRBANK, H. A. and LEE, D. M. 1957. The Fifth International Conference of Low Temperatures Physics and Chemistry, Wisconsin.
FAIRBANK, H. A. and WILKS, J. 1955. Proc. Roy. Soc. (London), A, **231**, 545.
GOODMAN, B. B. 1953. Proc. Phys. Soc. (London), A, **66**, 217.
GORTER, C. J., TACONIS, K. W., and BEENAKKER, J. J. M. 1951. Physica, **17**, 841.
KAPITZA, P. L. 1941. J. Phys. (U.S.S.R.), **4**, 181.
KHALATNIKOV, I. M. 1952. J. Exptl. Theoret. Phys. (U.S.S.R.), **22**, 687.
KOLSKY, H. 1953. Stress waves in solids (Oxford University Press, London).
KRONIG, R. and THELLUNG, A. 1950. Physica, **16**, 678.
MIEDEMA, A. R., POSTMA, H., and STEENLAND, M. 1958. Kamerlingh Onnes Conference of Low Temperature Physics.
RAYLEIGH, LORD. 1894. Theory of sound (Macmillan & Co., Ltd., London).
STRANDBERG, M. W. P. 1958. Phys. Rev. **110**, 65.
WILSON, A. H. 1953. The theory of metals (Cambridge University Press, London).

SOLID HARMONICS AS BASIS FUNCTIONS FOR CUBIC CRYSTALS[1]

D. D. BETTS[2]

ABSTRACT

The various sets of basis functions useful in discussing cubic crystals must include sets of symmetrized combinations of powers of the co-ordinates orthogonalized over the cellular polyhedron. Such polynomials are here called solid harmonics. A study of the actual solid harmonics reveals the limitations of the spherical cell approximation. The solid harmonics can be used to develop a new method over the cellular polyhedron of the body-centered cubic lattice or of the face-centered cubic lattice.

1. INTRODUCTION

Phenomena in solid state physics often require for their description functions of several variables having the symmetry of a crystal lattice, or a symmetry determined by that of the lattice. Thus it is desirable to have available complete sets of simple basis functions having the appropriate translational and rotational symmetry. Often, however, one considers functions defined throughout only one cellular polyhedron surrounding a lattice point in configuration space (or throughout a Brillouin zone in reciprocal space). In this paper we present sets of basis functions appropriate for the expansion of functions having rotational symmetry corresponding to the various inducible representations of the complete cubic group and defined throughout cellular polyhedron appropriate to the simple cubic, face-centered cubic, and body-centered cubic lattices respectively. We call such functions solid harmonics.

In Section 2 the various types of basis functions commonly used are discussed. In Section 3 the notation and method of computation of the new solid harmonics are discussed and tables of these functions are displayed. In Section 4 is discussed the adequacy in view of the results of Section 3 of the common approximation of replacing a cellular polyhedron by a sphere of equal volume. In Section 5 various applications of the new harmonics are mentioned, and application to the problem of integration of functions of cubic symmetry is discussed in detail.

One of the main aims of this paper is to show that the usual procedure of factoring mathematical functions occurring in solid state physics into radially and angularly dependent parts has limitations not usually mentioned.

Throughout this paper we shall be concerned mainly with functions associated with cubic lattices.

2. TYPES OF BASIS FUNCTIONS

Often we wish to have a set of basis functions having both the periodicity

[1]Manuscript received October 20, 1958.
Contribution from the Division of Pure Physics, National Research Council, Ottawa, and the Department of Physics, University of Alberta, Edmonton.
Issued as N.R.C. No. 5083.
[2]Permanent address: Department of Physics, University of Alberta, Edmonton, Alberta.

of the lattice and a rotation symmetry about a lattice point determined by the rotational symmetry of the lattice. In group theoretic language such functions are said to belong to a representation of the space group of the lattice, a direct product of a translation group and a rotation group. In order to ensure completeness of the set the basis functions are usually chosen as the eigenfunctions of some differential operator. The commonest and simplest such functions are the *plane waves*, linear combinations of functions of the form $e^{i\mathbf{k}\cdot\mathbf{r}}$ having a specified symmetry. The plane waves are solutions of the equation

$$(\nabla^2+k^2)u(\mathbf{r}) = 0. \tag{1}$$

Consider, for example, a simple cubic lattice of period 2 with a lattice point at the origin of a Cartesian co-ordinate system with axes parallel to the basis vectors of the lattice. The functions

$$\begin{aligned} W_{lmn} = {} & \cos l\pi x \cos m\pi y \cos n\pi z + \cos m\pi x \cos n\pi y \cos l\pi z \\ & + \cos n\pi x \cos l\pi y \cos m\pi z + \cos l\pi x \cos n\pi y \cos m\pi z \\ & + \cos n\pi x \cos m\pi y \cos l\pi z + \cos m\pi x \cos l\pi y \cos n\pi z \qquad l, m, n = 0, 1, 2 \ldots \end{aligned}$$

are plane waves which obviously have the periodicity of the lattice. The W_{lmn} are also invariant under all operations which leave the cube bounded by $x = \pm 1$, $y = \pm 1$, and $z = \pm 1$ invariant. In other words the W_{lmn} transform according that one of the 10 irreducible representations of the cubic rotation group (Bouckaert, Smoluchowski, and Wigner 1936) labelled s in the spectroscopic notation used by Howarth and Jones (1952). If now a second simple cubic lattice is added to the first in such a way as to produce a body-centered lattice the complete set of plane waves corresponding to the s representation consists of those of the above W_{lmn} for which $l+m+n$ is even. It is not difficult to write down in this way the plane waves corresponding to any of the irreducible representations of the space group of any lattice. However, it is not our purpose to exhibit lists of these functions here.

Another set of possible basis functions are the Mathieu functions which are solutions of the more complicated equation

$$[\nabla^2+k^2+V(\mathbf{r})]u(\mathbf{r}) = 0 \tag{2}$$

where $V(\mathbf{r})$ has a sinusoidal variation such that its minima occur at lattice points. Mathieu functions have been discussed extensively by Slater (1953). For a detailed discussion on wave propagation in periodic structures see Brillouin (1946). For a review of methods in the one-electron theory of solids using various types of wave functions, see Reitz (1955).

As opposed to functions which are continuous throughout a lattice, such as plane waves, there are functions in the theory of solids which have a "band" structure and are continuous only throughout a unit polyhedron of the lattice. The best known examples of such functions are the energies, E, as a function of the momentum vector $\mathbf{k}$, in the one-electron theory of solids. Another such function is the angular frequency ω of the normal modes of lattice vibration of solids again as a function of wave vector $\mathbf{k}$. (See for

example Born and Huang (1954).) Both $E(\mathbf{k})$ and $\omega(\mathbf{k})$ are in general found to be smoothly varying functions of $\mathbf{k}$ throughout a Brillouin zone.

As it is often desirable to expand such functions in a complete infinite set of basis functions (perhaps keeping as an approximation only a finite number of terms), the question arises as to what is the most appropriate basis set. The symmetrized plane waves already are orthogonalized over a cellular polyhedron giving them an important advantage. However, the simplest set of linearly independent functions having rotational symmetry consists of the combinations of powers of k_x, k_y, and k_z (or x, y, and z) as long as translational symmetry is not required. Such functions are linearly independent but not orthogonal to one another throughout a cellular polyhedron. Apparently the only way to construct an orthogonal set of such functions is through the well-known Schmidt orthogonalization process. We shall call sets of polynomials orthogonal over three-dimensional solids solid harmonics. The solid harmonics can be regarded as those solutions of the differential equation $\nabla^2 u = 0$ which are regular at the origin.

3. CONSTRUCTION OF SOLID CUBIC HARMONICS

The remainder of this paper will be concerned with the three monatomic lattices, simple cubic, body-centered cubic, and face-centered cubic. Special attention will be devoted to the latter two because only they are realized in actual crystals. The cellular polyhedra for the three lattices mentioned are a cube, a truncated cube or truncated octahedron, and a rhombohedral dodecahedron respectively. The latter two figures are illustrated, for example, by Reitz (1955).

The problem is to pick out those combinations of the powers of x, y, and z which transform according to each of the 10 irreducible representations of the cubic (or octahedral) group O_h and then to use the Schmidt process to make the linearly independent polynomials orthogonal to one another over each of the above solids and, for comparison, a sphere.

Actually, for the cube it is unnecessary to use the Schmidt process because polynomials of the form

$$P_l(x)P_m(y)P_n(z) \qquad l, m, n = 0, 1, 2 \ldots,$$

where $P_l(x)$ is the Legendre polynomial of degree l, are orthogonal to one another already over the cube bounded by $x = \pm 1$, $y = \pm 1$, $z = \pm 1$. Thus the complete orthogonal set of polynomials in x, y, z over the cube belonging to representation s (invariant under any permutation or change of sign of x, y, and z) are of the form

$$\begin{aligned}(3)\qquad C_s^{2l+2m+2n}(x, y, z) = {} & P_{2l}(x)P_{2m}(y)P_{2n}(z)+P_{2m}(x)P_{2n}(y)P_{2l}(z) \\ & +P_{2n}(x)P_{2l}(y)P_{2m}(z)+P_{2l}(x)P_{2n}(y)P_{2m}(z)+P_{2n}(x)P_{2m}(y)P_{2l}(z) \\ & +P_{2m}(x)P_{2l}(y)P_{2n}(z) \qquad l, m, n = 0, 1, 2, \ldots.\end{aligned}$$

The general expressions for the solid cubic harmonics belonging to the other representations are also easily written down. Picking out the combinations of x, y, and z having the symmetry appropriate to a given representation

is accomplished with the aid of the group character tables of Bouckaert, Smoluchowski, and Wigner (1936) or more directly with the aid of the polynomials of Von der Lage and Bethe (1947). It must be remembered that the x, y, and z of the latter authors are really the direction cosines x/r, y/r, and z/r in our notation. For example, $1, x^2+y^2+z^2 = r^2$, $x^4+y^4+z^4$, r^4 are linearly independent polynomials belonging to the s representation while x, y, z, x^3, y^3, z^3, r^2x, r^2y, r^2z belong to the p representation.

Performing the integration necessary to obtain an orthogonal set is elementary and straightforward but, especially in the case of the truncated cube or truncube, extremely tedious. All the integrals needed in order to construct the orthonormal sets of all truncubic, dodecahedral, and spherical cubic harmonics of degree less than five are listed in the Appendix.

It is important to have an adequate notation for the solid harmonics. The kernel or principal letter of the symbol indicates the region over which the harmonics are orthogonal, C, S, T, and D corresponding respectively to cube, sphere, truncube, and dodecahedron. The representation (giving the transformation properties of the harmonics) is indicated by a spectroscopic subscript. The harmonics are further distinguished by two numerical superscripts whose sum indicates the degree of the harmonic. Thus T_p^{21} indicates one of the harmonics of degree three with one nodal plane through the origin normalized and orthogonal to all other p type harmonics over the unit truncube. The order in which the independent polynomials are taken to form successive harmonics of the same representation is determined first by degree; then an arbitrary but consistent choice is made among those of the same degree. All solid harmonics of degree less than five are listed in Table I.

It should be noticed that the solid spherical harmonics S_a^{mn} orthogonal over the unit sphere are just products of Von der Lage and Bethe's (1947) Kubic Harmonics and simple polynomials in r.

4. ADEQUACY OF THE APPROXIMATION OF SPHERICAL CELLULAR POLYHEDRA

In the cellular method of Wigner and Seitz (1933) in the one-electron theory of solids the approximation is made of treating the cellular polyhedra as spheres. Only in a "spherical cellular polyhedron" can the one-electron Schrödinger equation be separated into radially and angularly dependent equations. If the usual assumption is made that the potential depends on r only then the solutions of the angularly dependent equation are the ordinary surface spherical harmonics, $Y_{lm}(\theta, \varphi)$. The approximation of a spherical cell becomes more realistic if the combinations of the Y_{lm} used transform according to one of the irreducible representations of the cubic or octahedral group O_h. Such combinations are the Kubic Harmonics of Von der Lage and Bethe (1947). In one sense the same approximation has been used by Houston (1948) in developing his method for determining the specific heat of lattices from the frequency distribution of the normal modes of the lattice vibrations. It must be recognized that various authors such as Horton and Schiff (1958) and Howarth and Jones (1952) have compensated in a reasonable way for the assumption of a spherical cell.

TABLE I

Solid harmonics orthonormal over polyhedra of cubic symmetry

Irreducible representation of cubic group	Sphere	Truncated cube	Rhombohedral dodecahedron
s	$S_s^0 = 1$	$T_s^0 = 1$	$D_s^0 = 1$
	$S_s^2 = \frac{5(7)^{\frac{1}{2}}}{2(3)^{\frac{1}{2}}}\left[r^2-\frac{3}{5}\right]$	$T_s^2 = \frac{32(15)^{\frac{1}{2}}}{(1129)^{\frac{1}{2}}}\left[r^2-\frac{19}{32}\right]$	$D_s^2 = \frac{8(15)^{\frac{1}{2}}}{(29)^{\frac{1}{2}}}\left[r^2-\frac{3}{8}\right]$
	$S_s^4 = \frac{63(11)^{\frac{1}{2}}}{(62)^{\frac{1}{2}}}\left[r^4-\frac{10}{9}r^2+\frac{5}{21}\right]$	$T_s^4 = \frac{420(6774)^{\frac{1}{2}}}{(20,636,963)^{\frac{1}{2}}}\left[r^4-\frac{51,305}{47,418}r^2+\frac{34,201}{158,060}\right]$	$D_s^4 = \frac{3360(29)^{\frac{1}{2}}}{(656,113)^{\frac{1}{2}}}\left[r^4-\frac{593}{812}r^2+\frac{10,039}{97,440}\right]$
	$S_s^{04} = \frac{5(77)^{\frac{1}{2}}}{4}\left[x^4+y^4+z^4-\frac{3}{5}r^4\right]$	$T_s^{04} = \frac{160(123,821,778)^{\frac{1}{2}}}{(21,323,916,097)^{\frac{1}{2}}}\left[x^4+y^4+z^4 -\frac{109,377,783}{165,095,704}r^4-\frac{21,327,933}{330,191,408}r^2 +\frac{277,157,435,631}{6,523,756,743,560}\right]$	$D_s^{04} = \frac{60(4,592,791)^{\frac{1}{2}}}{(8,042,273)^{\frac{1}{2}}}\left[x^4+y^4+z^4 -\frac{883,605}{656,113}r^4-\frac{305,116}{656,113}r^2-\frac{1,491,281}{26,244,520}\right]$
p (triply degenerate)	$S_p^1 = (5)^{\frac{1}{2}}z$	$T_p^1 = \frac{4(6)^{\frac{1}{2}}}{(19)^{\frac{1}{2}}}z$	$D_p^1 = 2(2)^{\frac{1}{2}}z$
	$S_p^{21} = \frac{21}{2}\left[r^2z-\frac{5}{7}z\right]$	$T_p^{21} = \frac{120(266)^{\frac{1}{2}}}{(34,201)^{\frac{1}{2}}}\left[r^2z-\frac{409}{570}z\right]$	$D_p^2 = \frac{720(7)^{\frac{1}{2}}}{(10,039)^{\frac{1}{2}}}\left[r^2z-\frac{41}{90}z\right]$
	$S_p^3 = \frac{5(21)^{\frac{1}{2}}}{2}\left[z^3-\frac{3}{5}r^2z\right]$	$T_p^3 = \frac{960(9,097,466)^{\frac{1}{2}}}{(39,971,472,229)^{\frac{1}{2}}}\left[z^3-\frac{234,989}{273,603}r^2 +\frac{242,707}{1,641,648}z\right]$	$D_p^3 = \frac{8(421,638)^{\frac{1}{2}}}{(65,833)^{\frac{1}{2}}}\left[z^3-\frac{7983}{10,039}r^2z+\frac{625}{10,039}z\right]$

(and harmonics obtained from above by replacing z by x or by y)

TABLE I (*continued*)

Solid harmonics orthonormal over polyhedra of cubic symmetry

Irreducible representation of cubic group	Sphere	Truncated cube	Rhombohedral dodecahedron
d^1 (doubly degenerate)	$S_{d^1}^{2(-)} = \frac{(35)^{\frac{1}{2}}}{2}[x^2-y^2]$	$T_{d^1}^{2(-)} = \frac{6(10)^{\frac{1}{2}}}{7}[x^2-y^2]$	$D_{d^1}^{2(-)} = 3(2)^{\frac{1}{2}}[x^2-y^2]$
	$S_{d^1}^{2(+)} = \frac{(35)^{\frac{1}{2}}}{(12)^{\frac{1}{2}}}[2z^2-x^2-y^2]$	$T_{d^1}^{2(+)} = \frac{2(30)^{\frac{1}{2}}}{7}[2z^2-x^2-y^2]$	$D_{d^1}^{2(+)} = (6)^{\frac{1}{2}}[x^2-y^2]$
	$S_{d^1}^{22(-)} = \frac{9(55)^{\frac{1}{2}}}{4}\left[r^2(x^2-y^2)-\frac{7}{9}(x^2-y^2)\right]$	$T_{d^1}^{22(-)} = \frac{564{,}480(2)^{\frac{1}{2}}}{(3{,}440{,}637{,}523)^{\frac{1}{2}}}\left[r^2(x^2-y^2)-\frac{108{,}131}{131{,}712}(x^2-y^2)\right]$	$D_{d^1}^{22(-)} = \frac{10{,}080(2)^{\frac{1}{2}}}{(423{,}743)^{\frac{1}{2}}}\left[r^2(x^2-y^2)-\frac{355}{672}(x^2-y^2)\right]$
	$S_{d^1}^{22(+)} = \frac{3(165)^{\frac{1}{2}}}{4}\left[r^2(2z^2-x^2-y^2)-\frac{7}{9}(2z^2-x^2-y^2)\right]$	$T_{d^1}^{22(+)} = \frac{188{,}160(6)^{\frac{1}{2}}}{(3{,}440{,}637{,}523)^{\frac{1}{2}}}\left[r^2(2z^2-x^2-y^2)-\frac{108{,}131}{131{,}712}(2z^2-x^2-y^2)\right]$	$D_{d^1}^{22(+)} = \frac{3360(6)^{\frac{1}{2}}}{(423{,}743)^{\frac{1}{2}}}\left[r^2(2z^2-x^2-y^2)-\frac{355}{672}(2z^2-x^2-y^2)\right]$
	$S_{d^1}^{4(-)} = \frac{7(165)^{\frac{1}{2}}}{4}\left[(x^4-y^4)-\frac{6}{7}r^2(x^2-y^2)\right]$	$T_{d^1}^{4(-)} = \frac{240(24{,}084{,}462{,}661)^{\frac{1}{2}}}{(3{,}256{,}830{,}958{,}481)^{\frac{1}{2}}}\left[(x^4-y^4)-\frac{2{,}867{,}072{,}706}{3{,}440{,}637{,}523}r^2(x^2-y^2)-\frac{266{,}279{,}985}{6{,}881{,}275{,}046}(x^2-y^2)\right]$	$D_{d^1}^{4(-)} = \frac{40(8{,}898{,}603)^{\frac{1}{2}}}{(2{,}677{,}049)^{\frac{1}{2}}}\left[(x^4-y^4)-\frac{440{,}403}{423{,}743}r^2(x^2-y^2)+\frac{26{,}457}{423{,}743}(x^2-y^2)\right]$
	$S_{d^1}^{4(+)} = \frac{7(55)^{\frac{1}{2}}}{4}\left[(2z^4-x^4-y^4)-\frac{6}{7}r^2(2z^2-x^2-y^2)\right]$	$T_{d^1}^{4(+)} = \frac{80(72{,}253{,}387{,}983)^{\frac{1}{2}}}{(3{,}256{,}830{,}958{,}481)^{\frac{1}{2}}}\left[(2z^4-x^4-y^4)-\frac{2{,}867{,}072{,}706}{3{,}440{,}637{,}523}r^2(2z^2-x^2-y^2)-\frac{266{,}279{,}985}{6{,}881{,}275{,}046}(2z^2-x^2-y^2)\right]$	$D_{d^1}^{4(+)} = \frac{40(2{,}966{,}201)^{\frac{1}{2}}}{(2{,}677{,}049)^{\frac{1}{2}}}\left[(2z^2-x^2-y^2)-\frac{440{,}403}{423{,}743}r^2(2z^2-x^2-y^2)+\frac{26{,}457}{423{,}743}(2z^2-x^2-y^2)\right]$

TABLE I (*concluded*)

Solid harmonics orthonormal over polyhedra of cubic symmetry

Irreducible representation of cubic group	Sphere	Truncated cube	Rhombohedral dodecahedron
d^2 (triply degenerate)	$S^2_{d^2} = (35)^{\frac{1}{2}}xy$	$T^2_{d^2} = \frac{24(5)^{\frac{1}{2}}}{(71)^{\frac{1}{2}}}xy$	$D^2_{d^2} = \frac{12(5)^{\frac{1}{2}}}{(7)^{\frac{1}{2}}}$
	$S^{22}_{d^2} = \frac{9(55)^{\frac{1}{2}}}{2}\left[xyr^2 - \frac{7}{9}xy\right]$	$T^{22}_{d^2} = \frac{2304(497)^{\frac{1}{2}}}{(4{,}867{,}621)^{\frac{1}{2}}}\left[xyr^2 - \frac{6065}{6816}xy\right]$	$D^{22}_{d^2} = \frac{20{,}160(7)^{\frac{1}{2}}}{(395{,}347)^{\frac{1}{2}}}\left[xyr^2 - \frac{1115}{2352}xy\right]$
	$S^4_{d^2} = \frac{7(165)^{\frac{1}{2}}}{2}\left[xyz^2 - \frac{1}{7}xyr^2\right]$	$T^4_{d^2} = \frac{2400(34{,}073{,}347)^{\frac{1}{2}}}{(56{,}419{,}187{,}959)^{\frac{1}{2}}}\left[xyz^2 + \frac{2{,}095{,}619}{121{,}690{,}525}xyr^2 - \frac{1{,}943{,}277}{24{,}338{,}105}xy\right]$	$D^4_{d^2} = \frac{8(41{,}511{,}435)^{\frac{1}{2}}}{(170{,}063)^{\frac{1}{2}}}\left[xyz^2 - \frac{90{,}529}{395{,}347}xyr^2 + \frac{11{,}820}{395{,}347}xy\right]$
f^1	$S^3_{f^1} = 3(35)^{\frac{1}{2}}xyz$	$T^3_{f^1} = \frac{96(210)^{\frac{1}{2}}}{(3199)^{\frac{1}{2}}}xyz$	$D^3_{f^1} = \frac{48(21)^{\frac{1}{2}}}{(37)^{\frac{1}{2}}}xyz$
f^2 (triply degenerate)	$S^3_{f^2} = \frac{3}{2}(35)^{\frac{1}{2}}z(x^2-y^2)$	$T^3_{f^2} = \frac{96(105)^{\frac{1}{2}}}{(12929)^{\frac{1}{2}}}z(x^2-y^2)$	$D^3_{f^2} = 12(3)^{\frac{1}{2}}z(x^2-y^2)$
g	$S^4_g = \frac{(1155)^{\frac{1}{2}}}{2}xy(x^2-y^2)$	$T^4_g = \frac{16(105)^{\frac{1}{2}}}{(1057)^{\frac{1}{2}}}xy(x^2-y^2)$	$D^4_g = \frac{40(21)^{\frac{1}{2}}}{(19)^{\frac{1}{2}}}xy(x^2-y^2)$

(All harmonics of the other three types are of degree greater than four)

An examination of the integrals in the Appendix shows that the average values of the even powers of r over the unit sphere, $\langle r^{2n}\rangle_s$ are very nearly equal to the average values over the unit truncube, $\langle r^{2n}\rangle_t$. The greatest difference, amounting to 4%, is found for $\langle r^8\rangle$. We take this as some evidence that for functions of r only the approximation of replacing the cellular polyhedron of a body-centered lattice (or the Brillouin zone of a face-centered lattice) by a sphere is not a bad approximation. If we examine the table of harmonics, we notice that the corresponding coefficients of the S_s and the T_s are quite similar as indicated by a comparison of the unnormalized harmonics below.

$$S_s^2 \sim r^2-.600 \qquad T_s^2 \sim r^2-.594$$

$$S_s^4 \sim r^4-1.111\,r^2+.238 \qquad T_s^4 \sim r^4-1.085\,r^2+.216$$

$$S_s^{04} \sim x^4+y^4+z^4-.600\,r^4 \qquad T_s^{04} \sim x^4+y^4+z^4-.663\,r^4-.064\,r^2+.042$$

Also the T_{d^1} and S_{d^1} are rather similar.

$$S_{d^1}^{22} \sim r^2-.778 \qquad T_{d^1}^{22} \sim r^2-.821$$

$$S_{d^1}^4 \sim x^2+y^2-.857\,r^2 \qquad T_{d^1}^{22} \sim x^2+y^2-.834\,r^2-.039$$

Thus we tentatively propose on the basis of fairly meagre evidence that the spherical cell may be a fairly good approximation to the cellular polyhedron of the body-centered cubic lattice for s and d^1 type functions. For p and d^2 type functions, however, Table I seems to indicate that the spherical approximation is inadequate. Also in the face-centered lattice the spherical approximation seems to be poor for functions of any of the types s, p, d^1, and d^2. No conclusions can be drawn about functions of the other six types. Nevertheless, it is often expedient to use the spherical cell just because of the great computational complexity involved in solving non-separable three-dimensional Schrödinger equations.

5. APPLICATIONS

The solid harmonics should be useful as basis functions for expansion of smoothly varying functions of cubic symmetry in cellular polyhedra or Brillouin zones of cubic lattices. However, the simple unorthogonalized polynomials are adequate in many cases, such as in the expansion of the one-electron energy, $E(\mathbf{k})$, as

$$E(\mathbf{k}) = E_0+ak^2+bk^4+c(k_x^4+k_y^4+k_z^4)+\ldots\,. \tag{4}$$

Only when *integration* of the expanded function is to be performed is it more useful to expand in terms of the orthogonal solid harmonics.

In the theory of specific heats of lattices, for example, it is necessary to integrate over the cellular polyhedron a function of $\mathbf{k}$ in general so complicated that one can obtain analytic expressions for it only for certain special sets of values of $\mathbf{k}$, typically $k_x = 0$ or $k_x = k_y$. Houston (1948) took advantage of this fact in developing his method, which samples the integrand in the

three directions $k_z = k_y = 0$, $k_x = k_y$ and $k_z = 0$, and $k_x = k_y = k_z$ and obtains the angular dependence of the integrand by fitting a linear combination of the first three Kubic Harmonics. Houston's method suffers from the disadvantage that in converting the variable of integration from k to ω infinities are artificially introduced. This difficulty is avoided by Horton and Schiff (1958), but they still use the Kubic Harmonics as interpolative functions.

We propose that a use of the appropriate solid harmonics for such an interpolative scheme is ultimately capable of the greatest accuracy. The method can be described in general terms but we prefer to illustrate it with a concrete example. In Fig. 1 (*a*) is illustrated the 1/48th of the rhombohedral dodecahedron containing all the non-equivalent points. Following Bouckaert, Smoluchowski, and Wigner (1936) we have labelled the origin and the other

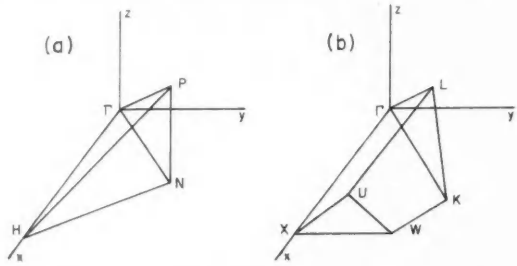


FIG. 1. (*a*) The polyhedron containing all the symmetrically non-equivalent points in the rhombohedral dodecahedron.
(*b*) The polyhedron containing all the symmetrically non-equivalent points of the truncated cube.

corner points Γ, H, N, and P respectively. The scale of our figure is conveniently fixed by letting H have co-ordinates $(1, 0, 0)$ then N $(\frac{1}{2}, \frac{1}{2}, 0)$ and P $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. Instead of the completely orthogonal set of solid harmonics of Table I we introduce the s-type harmonics

$$
\begin{aligned}
d^0 &= 1 \\
d^1 &= x^2+y^2+z^2-3/8 \\
d^2 &= x^4+y^4+z^4-9/80 \\
d^{11} &= x^2y^2+y^2z^2+z^2x^2-7/240 \qquad (5) \\
&\cdots\cdots\cdots \\
d^{lmn} &= x^{2l}y^{2m}z^{2n}+y^{2l}z^{2m}x^{2n}+z^{2l}x^{2m}y^{2n}+x^{2l}z^{2m}y^{2n}+y^{2l}x^{2m}z^{2n} \\
&\qquad +z^{2l}y^{2m}x^{2n}+a^{lmn}.
\end{aligned}
$$

The constant a^{lmn} is chosen to make the harmonic d^{lmn} orthogonal to d^0. The integrals in the Appendix are sufficient to enable us to write down immediately the first 11 harmonics of this kind.

Now let $f(x, y, z)$ be an s-type function which it is required to integrate over the rhombohedral dodecahedron. The expansion

$$f(x, y, z) = \sum_{l,m,n} b_{lmn}\, d^{lmn}(x, y, z) \tag{6}$$

is exact if the whole infinite set of d's is used. The expansion is of a certain degree of approximation k if all harmonics of degree $2k$ or less are used. Now

$$\int_{\text{dodecahedron}} f(x, y, z) = b_0 \int_{\text{dodecahedron}} d^0 \tag{7}$$

$= b_0 \times$ volume of dodecahedron because of the orthogonal property of the d's.

The "second approximation" to b_0 is found by solving the four equations derived by evaluating the quantities in eq. 6 at the four special corner points Γ, H, N, and P and keeping only the first four harmonics d^0, d^1, d^{11}, and d^2. The result is

$$b_0 = (71f_\Gamma + 9f_H + 336f_N - 56f_P)/360. \tag{8}$$

For the truncube we can define harmonics t^{lmn}. The integrals tabulated in the Appendix analogous to (5) allow us to write down the first 11 immediately, of which the first 4 are

$$\begin{aligned} t^0 &= 1 \\ t^1 &= x^2+y^2+z^2-19/32 \\ t^2 &= x^4+y^4+z^4-89/320 \\ t^{11} &= x^2y^2+y^2z^2+z^2x^2-71/960. \end{aligned} \tag{9}$$

A second degree approximation gives

$$b_0 = (3659f_\Gamma + 5037f_X - 3840f_K + 13{,}864f_L)/18{,}720 \tag{10}$$

for the truncube.

The integration formulas given above which sample the integrand only at four points are obviously quite inadequate for all but the smoothest functions. However, the extension of formulas like (8) and (10) to higher degrees of approximation is straightforward and should become quite accurate for a reasonably small number of points. Houston's method has been found quite adequate in most cases of practical interest by Betts, Bhatia, and Wyman (1956) when extended to include six points. By analogy for a three-dimensional rather than a two-dimensional region of integration perhaps of the order of 15 or 20 points would be quite adequate. Further investigation of such formulas will be carried out.

It appears that the same sort of formulas as (8) and (10) could be obtained using plane waves instead of solid harmonics. There are two objections to such a proposal. First, one would expect that an expansion of a smoothly

varying function such as ω in solid harmonics would converge more rapidly than an expansion in plane waves just as in one dimension we normally expand smoothly varying functions in power series rather than a more slowly converging Fourier series. Second, it turns out that it is very difficult to make a reasonable choice of points for integrand evaluation such that the equations like those derived from (6) are all linearly independent.

No obvious application of solid harmonics can be made to the problem of determining crystal wave functions because of the great complexity of non-separable three-dimensional one-electron Schrödinger equations.

6. CONCLUSIONS

Among the various types of basis functions used in the study of solids ought to be included the solid harmonics. It would be of considerable interest to have a method of generating solid harmonics less laborious than the Schmidt process, although perhaps an adequate number for many purposes can be obtained in this way. The spherical cell approximation seems to be fairly good for s and d^1 type functions in the body-centered cubic lattice unit cell. It seems to be quite poor for functions of p or d^2 symmetry in the above cell or for functions of s, p, d^1, or d^2 symmetry in the unit cell of the face-centered cubic lattice. Nothing can be concluded about the spherical approximation for functions of any of the other six symmetry types in either lattice.

The use of solid harmonics promises to provide a new method of numerical integration of smoothly varying functions of cubic symmetry which should require far fewer points than the usual straightforward method of numerical integration for the same degree of accuracy. For integrands with discontinuities, however, the usual method with its "unprejudiced" weights should be better.

ACKNOWLEDGMENT

The author gratefully acknowledges the hospitality of Dr. T. Y. Wu of the National Research Council during the summer of 1958.

REFERENCES

BETTS, D. D., BHATIA, A. B., and WYMAN, M. 1956. Phys. Rev. **104**, 37.
BOUCKAERT, L. P., SMOLUCHOWSKI, R., and WIGNER, E. P. 1936. Phys. Rev. **50**, 58.
BRILLOUIN, L. 1946. Wave propagation in periodic structures (McGraw-Hill Book Co., New York).
BORN, M. and HUANG, K. 1954. Dynamical theory of crystal lattices (The Clarendon Press, Oxford).
HORTON, G. K. and SCHIFF, H. 1958. To be published.
HOUSTON, W. V. 1948. Revs. Modern Phys. **20**, 161.
HOWARTH, D. J. and JONES, H. 1952. Proc. Phys. Soc. A, **65**, 355.
VON DER LAGE, F. C. and BETHE, H. A. 1947. Phys. Rev. **71**, 612.
REITZ, J. R. 1955. Solid state physics, Vol. 1 (Academic Press, Inc., New York), p. 1.
SLATER, J. C. 1953. Tech. Rept. No. 4, Solid State and Molecular Theory Group, Mass. Inst. Technol. (unpublished).
WIGNER, E. P. and SEITZ, F. 1933. Phys. Rev. **43**, 804.

APPENDIX

Normalized integrals of polynomials

Polynomial	Region of integration		
	Sphere	Truncube	Dodecahedron
1	1	1	1
r^2	3/5	19/32	3/8
x^4	3/35	89/960	3/80
x^2y^2	1/35	71/2880	7/720
r^4	3/7	409/960	41/240
x^6	1/21	1237/21,504	83/5376
x^4y^2	1/105	1/120	31/16,128
$x^2y^2z^2$	1/315	457/276,480	37/48,384
r^6	1/3	6703/20,160	463/5376
x^8	1/33	377/9216	31/3840
x^6y^2	1/231	2857/645,120	1/1792
x^4y^4	1/385	743/645,120	37/134,400
$x^4y^2z^2$	1/1155	2573/6,451,200	53/403,200
r^8	3/11	38,281/134,400	2119/44,800

THE PRESSURE-INDUCED ROTATIONAL ABSORPTION SPECTRUM OF HYDROGEN

I. A STUDY OF THE ABSORPTION INTENSITIES[1]

Z. J. KISS,[2] H. P. GUSH,[3] AND H. L. WELSH

ABSTRACT

The pressure-induced infrared spectrum of H_2 and mixtures of H_2 with N_2, He, Ne, A, Kr, and Xe was measured in the region 300–1400 cm^{-1} at total gas pressures up to 250 atm at 300° K and, where possible, at 195° K and 85° K. The spectrum shows greatly broadened S lines ($\Delta J = +2$) with half widths which decrease as the temperature is lowered. The integrated absorption coefficient of the band is of the form $\alpha_1\rho_{H_2}\rho_f$, where ρ_f is the density of the perturbing gas, except in the case of Xe for which a cubic term, $\alpha_2\rho_{H_2}\rho_f^2$, is also necessary. The binary coefficient α_1 increases by a factor of 28 in going from He to Xe. The theoretical band intensity, calculated on the basis of quadrupole interaction alone, is in good agreement with the experimental value only for Xe as perturbing gas; in other cases the calculated value is appreciably less than the observed value. The shape of the absorption contours suggests that the S lines are overlaid by a continuum increasing in intensity towards lower frequencies. This continuum is interpreted as the counterpart of the Q_R component in the vibrational band, that is, a collision-induced absorption due to overlap interaction in which the relative kinetic energy of the collision partners changes in the absorption process.

INTRODUCTION

The infrared vibrational spectrum of hydrogen, studied under widely different experimental conditions, has served to delineate many features of collision-induced spectra (Chisholm and Welsh 1954; Hare and Welsh 1958). A simple picture of the origin of these spectra is the following: The intermolecular forces operative during collisions produce a distortion of the electron distributions of the molecules. The resulting electric dipole moment of the system is modulated by the vibration and rotation of the colliding molecules, so that molecular frequencies, which are inactive because of symmetry in the free molecule, appear in infrared absorption. In the region of densities in which only binary collisions are important, the intensity of the induced absorption is proportional to the square of the density of the pure gas; the enhancement of the absorption by a foreign gas is proportional to the product of the densities of the absorbing gas and the perturbing gas. As shown by the structure of the fundamental vibrational band of hydrogen the selection rule for rotational transitions in homonuclear diatomic molecules is $\Delta J = 0, \pm 2$, where J is the total angular momentum quantum number; this selection rule is a consequence of the fact that these molecules have an electrical center of symmetry which coincides with the center of gravity.

[1]Manuscript received December 5, 1958.
Contribution from the McLennan Laboratory, University of Toronto, Toronto, Ontario. This research was supported by Grant No. T54 of the National Research Council of Canada.
[2]Holder of a scholarship from the National Research Council, 1957–58.
[3]Holder of a scholarship from the National Research Council, 1955–57. Present address: Laboratoire Aimé Cotton, C.N.R.S., Bellevue (S. et O.), France.

The theory of the induced fundamental band of hydrogen has been given by Van Kranendonk and Bird (1951), Van Kranendonk (1952, 1957, 1958), and Britton and Crawford (1958), who have shown that the absorption must be explained in terms of electron overlap and molecular quadrupole interactions. The overlap interaction appears to be almost spherically symmetric and thus produces mainly transitions for which $\Delta J = 0$ (Q branch); the quadrupole interaction on the other hand is strongly dependent on the mutual orientation of the molecules in a pair and produces transitions for which $\Delta J = \pm 2$ (S and O branches) and, in addition, a small contribution to the Q branch intensity.

Any homonuclear diatomic molecule should also show an induced spectrum in which purely rotational transitions are involved. However, only in the case of hydrogen would this spectrum lie in a region accessible to the more conventional infrared techniques. The Raman spectrum of hydrogen at room temperature shows four rotational lines, $S(J)$ with $J = 0, 1, 2, 3$, at 354.4, 587.1, 814.4, and 1034.7 cm^{-1}, respectively (Stoicheff 1957); in the induced infrared spectrum these lines are all in the region of CsBr optics. From the theory of the vibrational band one would expect that the rotational lines of hydrogen would appear in absorption mainly as a result of quadrupole interaction; it is therefore possible that, under certain conditions, the value of the quadrupole moment of the hydrogen molecule could be calculated from the intensity of the absorption.

The $S(2)$ and $S(3)$ rotational lines were observed by Ketelaar, Colpa, and Hooge (1955), and in a recent paper Colpa and Ketelaar (1958) have given data for the $S(1)$, $S(2)$, and $S(3)$ lines in pure hydrogen and in mixtures with helium, argon, nitrogen, and carbon dioxide at −60°, 25°, and 80° C. The present investigation was designed to obtain experimental data for the whole rotational spectrum in the pure gas and in mixtures with nitrogen, helium, neon, argon, krypton, and xenon at room temperature. The heavier rare gases were included since it seemed probable that for these the quadrupole effect would be relatively much larger than the overlap effect, as compared with perturbing molecules with smaller polarizabilities. Experiments on the vibrational band of hydrogen have shown that many features of the spectrum become clearer at low temperatures; the investigation of the rotational spectrum was therefore extended whenever possible to 195° K and 85° K. In the present paper the data on the integrated absorption intensities are discussed; in a future communication an analysis of the absorption profiles will be given.

EXPERIMENTAL

To record the whole rotational spectrum of hydrogen it was necessary to use windows of caesium bromide in the pressure tube. Since caesium bromide is relatively soft some difficulties were experienced from the plastic flow of the crystal under pressure. In the design finally adopted circular windows, 25 mm in diameter and 13 mm thick, were used with a free area of 10 mm × 5 mm; the rectangular free area was used in preference to a circular area of the same size, since the shear strength for the former is greater than for

the latter as a consequence of the greater shear area. The deformation of the window under pressure was studied by placing a glass rod connected to an optical lever in contact with the outside surface of the window. The amount of extrusion as a function of the pressure is shown in Fig. 1 (*a*), the pressure being applied in each case for 10 minutes. The extrusion as a function of time is shown in Fig. 1 (*b*) for two different pressures. It was noted that the plastic flow takes place in a preferential direction in the crystal, so that additional

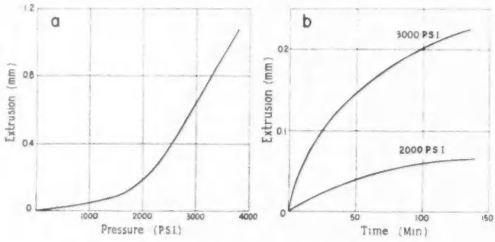


FIG. 1. Measurements of the extrusion of CsBr windows under pressure: (*a*) amount of extrusion as a function of pressure applied for 10-minute intervals, (*b*) extrusion as a function of time.

strength in the window might be obtained by a proper orientation of the crystal axes; however, no attempt was made to do this in this investigation. As a result of these experiments it was concluded that the windows used could support a pressure of 300 atm for an extended period without serious deformation. It was necessary, however, to renew the windows frequently since they had a tendency to crack when the pressure was released.

The high pressure absorption cell (Fig. 2) was of the transmission type with a path length of 26.7 cm and was designed to withstand pressures up to 500 atm. The body of the cell, A, was a stainless steel cylinder with a central bore, 1.2 cm in diameter, into which was inserted a light guide, B,

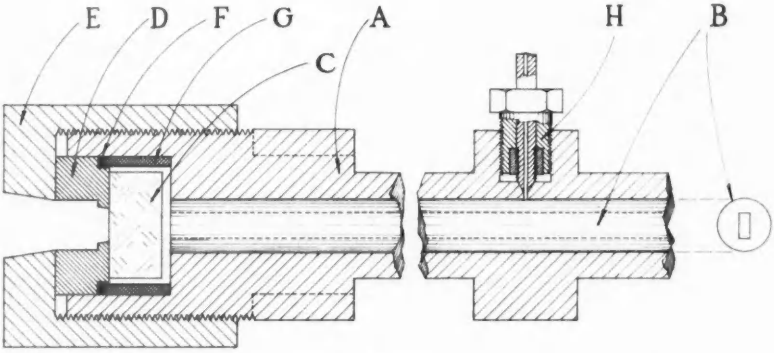


FIG. 2. Design of the absorption cell for high pressures and low temperatures.

with a rectangular cross section, 7 mm × 2.5 mm. The apertures of the cell were designed so that $f/4$ cones of radiation could be used to focus the source and the spectrometer slit on the entrance and exit ends of the light guide respectively. The window, C, was carried by the hardened steel ring, D, to which it was attached by a thin layer of rubber cement. The ring, D, was sealed to the body of the cell with pressure applied by the closing nut, E, to the lead gasket, F. The incorporation of a tightly fitting invar ring, G, in the design maintained the pressure seal at temperatures down to 80° K. In the low temperature experiments the tube was immersed in a trough containing liquid air or solid carbon dioxide and acetone, and insulated with styrofoam; CsBr lenses attached to the ends of the absorption cell by evacuated monel tubes were used as condensers.

Since some of the absorptions were very weak in the pressure range which could be used, a second absorption cell with a path length of 245 cm was also employed. The design of the end seals of this tube was the same as that described above, and a silver wave guide served as the light guide.

The spectra were recorded with a Perkin–Elmer 12C spectrometer equipped with a CsBr window and a thermocouple detector. In preliminary experiments with the cell at liquid air temperature and the chopper in its normal position in front of the spectrometer slit, it was noted that a negative signal was recorded in regions of low absorption. It was apparent that the thermocouple was radiating energy to the cold cell and that this chopped back-radiation was responsible for the negative signal. The effect was eliminated by placing the chopper between the source and the cell. In the experiments at room temperature the source was a Globar operated at 4 amp from a Sorenson voltage regulator. Better results were obtained in the later experiments at low temperatures using a commercial Welsbach mantle heated by an oxygen-gas flame. Since the intensity of both of these sources falls off rapidly at long wave lengths it was found necessary to alter the slit width of the spectrometer and the gain of the amplifier for each of the four regions 5–12 μ, 12–18 μ, 18–26 μ, and 26–33 μ; however, an effective resolution of at least 5 cm^{-1} was preserved throughout the range. To reduce the effects of stray radiation the metal plate of the beam chopper was replaced by a fused quartz plate so that only radiation of wave length greater than 4 μ was detected. With this device the effect of stray radiation was negligible except at wave lengths greater than 25 μ; for this region the line of infinite absorption on the recorder trace was obtained by covering the slit of the spectrometer with a thick plate of rock salt. The frequency calibration was obtained from the atmospheric water vapor and carbon dioxide bands in the region 5–9 μ and 23–33 μ, supplemented in the region 9–11 μ by the spectrum of ammonia gas at a pressure of 10 mm Hg in an absorption cell 10 cm long.

The gases were compressed with a hand-operated 1500-atm gas compressor, and the pressure in the cell was measured with a calibrated Bourdon tube gauge. To eliminate water vapor absorption the gases were dried by passage through a cold trap. Mixtures of gases were made at lower pressures in the chamber of the gas compressor and it was assumed that the density

of each gas could be calculated from its partial pressure. The densities of the gases were calculated from the measured pressures and temperatures; isothermal data for all the gases are available in the literature for the temperature and pressure ranges used.

In experiments with compressed hydrogen gas at lower temperatures ortho–para conversion, catalyzed by the steel of the absorption cell, takes place slowly. Since it is essential to know the ortho–para ratio in calculating the absorption coefficients, the experimental data were obtained under conditions for which there was no conversion (normal hydrogen) or for which there was complete conversion (equilibrium hydrogen for the temperature in question). To obtain normal hydrogen at lower temperatures a leak was introduced in the absorption cell so that the contents of the cell were renewed from the chamber of the gas compressor at room temperature every 3 minutes; control experiments showed that in this period of time the ortho–para conversion was negligible. An equilibrium ortho–para mixture at 85° K was prepared by passing hydrogen through a copper coil filled with alumina gel coated with chromic oxide and cooled to 85° K; some of the catalyst was also placed in the high pressure cell itself. The conversion was followed by recording the signal level at the position of the $S(1)$ line and appeared to be complete after 3 hours.

EXPERIMENTAL RESULTS

Table I summarizes the experimental conditions for which one or more sets of absorption profiles of the rotational band were obtained. Each experiment,

TABLE I

Outline of the experiments

Perturbing gas	Temperature, ° K	Optical path length, cm	Maximum density, Amgt
Normal H_2	300	27	200
	300	245	45
	195	27	100
	85	27	120
Equilibrium H_2	85	27	120
He	300	27	150
	300	245	70
	195	27	250
Ne	300	27	80
	300	245	50
	195	27	80
A	300	27	120
	195	27	120
N_2	300	27	120
	195	27	120
Kr	300	27	75
Xe	300	27	70

for a given perturbing gas and a given temperature, consisted in recording the spectrum from approximately 300 cm^{-1}, the low frequency limit of observation, to 1400 cm^{-1} for a series of six or more different pressures.

Examples of the absorption profiles ($\log_{10}(I_0/I)$ vs. frequency in cm^{-1}) are shown in Figs. 3, 4, 5, and 6. In these graphs the frequencies of the Raman lines of gaseous hydrogen are marked on the abscissa axis. In the band for

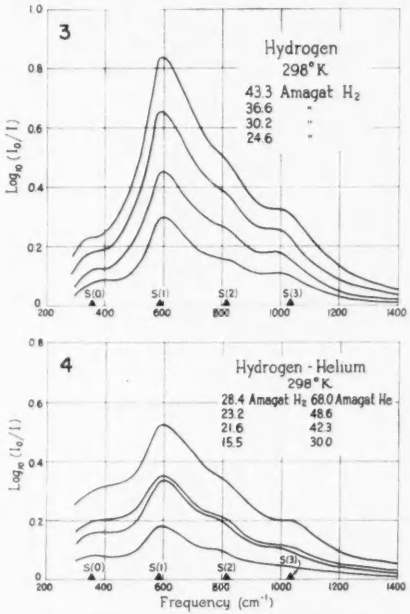


FIG. 3. Absorption of pure hydrogen in the region 300–1400 cm^{-1} at room temperature. Path length, 27 cm. The Raman frequencies of the gas are marked on the abscissa axis.

FIG. 4. Absorption of hydrogen–helium mixtures in the region 300–1400 cm^{-1} at room temperature. Path length, 245 cm.

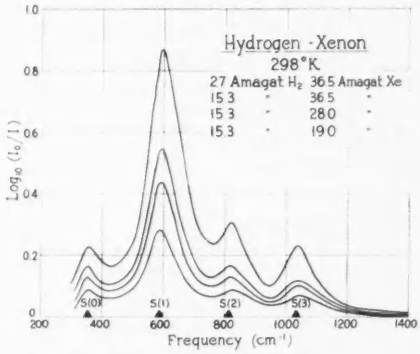


FIG. 5. Absorption of hydrogen–xenon mixtures in the region 300–1400 cm^{-1} at room temperature. Path length, 27 cm.

normal hydrogen at room temperature (Fig. 3) the four lines $S(0)$, $S(1)$, $S(2)$, and $S(3)$ are readily distinguished, although they are very broad and there is appreciable overlapping. For the H_2–Xe mixture (Fig. 5) the lines are comparatively sharp and it is evident that the maxima lie close to the Raman frequencies. In contrast, the H_2–He mixture (Fig. 4) shows a much weaker over-all absorption and the S lines are poorly defined.

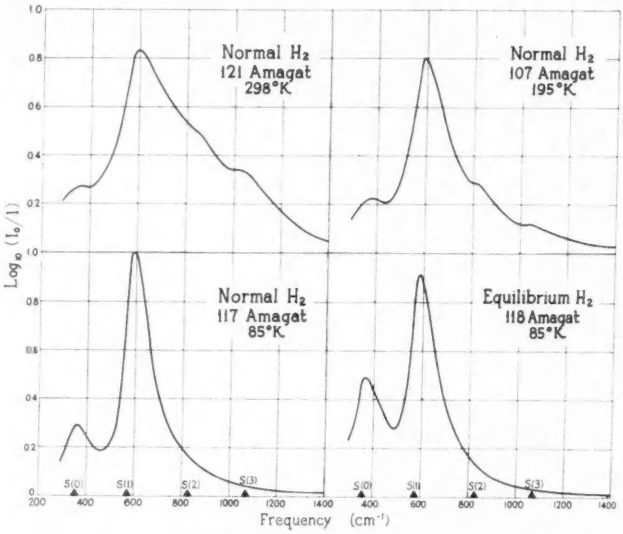


FIG. 6. Comparison of the absorptions of normal hydrogen at 298° K, 195° K, and 85° K and equilibrium hydrogen at 85° K. Path length, 27 cm.

The dependence of the absorption of pure hydrogen on temperature and ortho–para ratio is shown in Fig. 6. The half width of the lines decreases markedly with temperature as a consequence of the longer duration of collisions at lower temperatures. Only the $S(0)$ and $S(1)$ lines are recorded at 85° K since the higher rotational levels are not appreciably populated. The greater intensity ratio, $S(0)/S(1)$, for the equilibrium mixture at 85° K is due to the greater para concentration, 45.8%, as compared with 25% for normal hydrogen.

Integrated absorption intensities, $\int \alpha d\nu$, were obtained by multiplying the areas under the absorption profiles by $2.30/l$, where l is the path length. Since a small portion of the band at frequencies less than 300 cm^{-1} was not recorded, it was necessary to extrapolate the experimental curves by assuming a shape for the low frequency wing of the $S(0)$ line similar to that observed for the $S(1)$ line; the correction is quite small so that this procedure can be considered adequate. In Fig. 7 the values of $(1/\rho_{H_2})\int \alpha d\nu$ at 300° K are plotted against the density of the perturbing gas in Amagat units. In the case of gas mixtures $\int \alpha d\nu$ refers to the enhancement by the foreign gas and is obtained

by subtracting from the total intensity the intensity due to H_2–H_2 collisions, which can be calculated from the data for the pure gas. In all cases except the H_2–Xe mixture, $(1/\rho_{H_2})\int \alpha d\nu$ is a linear function of ρ_f, so that the integrated intensity has the form $\int \alpha d\nu = \alpha_1 \rho_{H_2}^2$ for the pure gas, and $\int \alpha d\nu = \alpha_1 \rho_{H_2} \rho_f$ for the mixtures. The constant, α_1, is thus a specific absorption coefficient for binary collisions and can be conveniently expressed in the units, cm^{-2} $Amgt^{-2}$, using the Amagat unit of density.

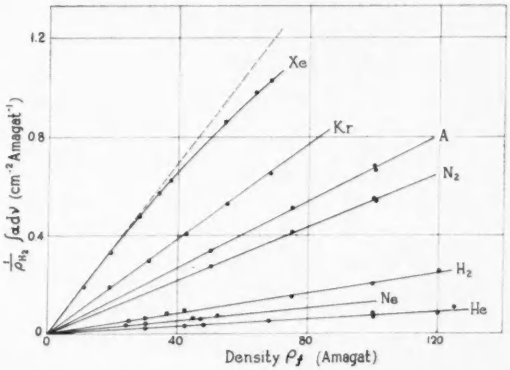


FIG. 7. The variation of the integrated absorption intensity of hydrogen in the region 300–1400 cm^{-1} with the density, ρ_f, of the perturbing gas.

The graph for the H_2–Xe mixture in Fig. 7 departs noticeably from a straight line; here the integrated intensity follows the relation,

$$\int \alpha d\nu = \alpha_1 \rho_{H_2} \rho_{Xe} + \alpha_2 \rho_{H_2} \rho_{Xe}^2$$

where $\alpha_1 = 1.74 \times 10^{-2}$ cm^{-2} $Amgt^{-2}$ and $\alpha_2 = -5.5 \times 10^{-5}$ cm^{-2} $Amgt^{-3}$. The small cubic term shows some effect of ternary collisions of the type Xe–H_2–Xe. According to general calculations of Van Kranendonk (1957) this term is given by the superposition of three effects: (*a*) the density dependence of the distribution function for pairs of molecules (finite volume effect), (*b*) the partial cancellation of the two moments induced in a ternary collision of the above type, and (*c*) non-additivity of the dipoles induced in ternary collisions. The cancellation effect (*b*) leads in general to a negative cubic term and is probably chiefly responsible for the cubic coefficient observed for the H_2–Xe mixture. The reason for the appearance of the cubic term in the single case of the H_2–Xe mixture may be due to the fact that the experiment was carried out close to the critical point (16° C) of xenon; it may therefore be indicative of molecular clustering in the neighborhood of the critical point. This hypothesis could be readily tested by extending the experiments to higher temperatures.

The effect of lower temperatures on the integrated intensity is illustrated in Fig. 8 where the results for normal hydrogen at 300°, 195°, and 85° K are

given. The decrease in the value of α_1 as the temperature is lowered is qualitatively explained by the greater distance of closest approach of the molecules in collision, giving a smaller average interaction. Figure 9 shows that increasing the para concentration at 85° K gives a slight decrease in α_1.

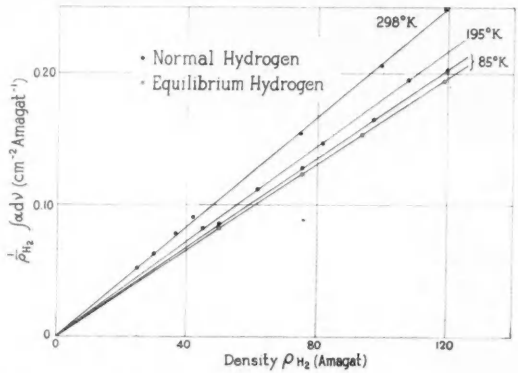


FIG. 8. The effect of temperature and ortho–para ratio on the integrated absorption intensity of hydrogen.

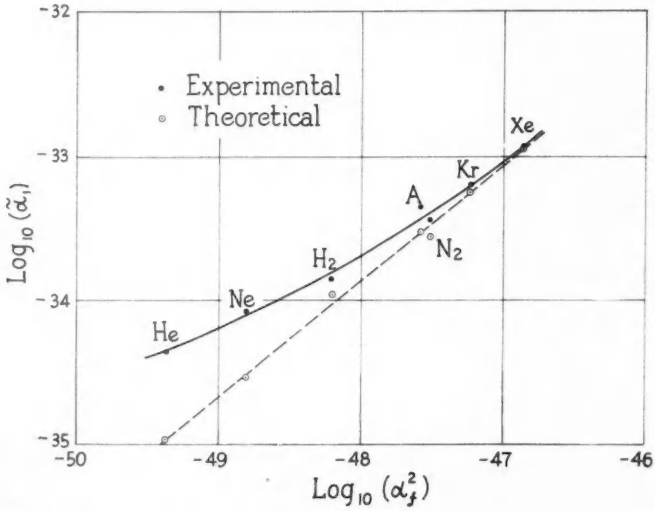


FIG. 9. Variation of the experimental values of $\tilde{\alpha}_1$ and calculated values of $\tilde{\alpha}_1$(quad.) with the molecular polarizability, α, of the perturbing gas.

For comparison with theory it is convenient to introduce a different measure of the intensity, $\tilde{\alpha}_1$, in which the effects of (*a*) the "apparent" induced absorption and (*b*) the varying absorption frequency, ν, have been eliminated. Apparent induced absorption is the slight enhancement of any absorption which

results from the long-range dipole–dipole interaction between molecules. According to Van Kranendonk (1957) experimental intensities can be corrected for apparent induced absorption by multiplication with the factor,

$$f(\rho) = \left(1+\frac{8\pi}{3}\alpha n\right)^{1/2}\left(1-\frac{4\pi}{3}\alpha n\right)^{3/2},$$

where α is the average electronic polarizability of the molecule and n is the number of molecules per cubic centimeter. In the most extreme case considered here this correction is of the order of 1%. The experimental absorption intensity contains the usual frequency factor, ν, the value of which varies over a wide range in the case of the rotational band. It is expedient to remove this frequency dependence by means of an average absorption frequency, $\bar{\nu}$, defined by the equation,

$$\bar{\nu} = \int \alpha d\nu / \int (\alpha/\nu) d\nu.$$

The frequency-corrected binary coefficients, $\tilde{\alpha}_1 = \alpha_1/\bar{\nu}$, are thus in the nature of transition probabilities; they will be expressed here in the form, cycles per second per molecule per cm^3 of hydrogen per molecule per cm^3 of perturbing gas, and have therefore the dimensions sec^{-1} cm^6. The essential results of the reduction of the experimental data are thus α_1, $\bar{\nu}$, and $\tilde{\alpha}_1$; the values of these are given in columns 3, 4, and 5 of Table III below.

CALCULATION OF THE BINARY ABSORPTION COEFFICIENTS DUE TO QUADRUPOLE INTERACTION

As shown by Van Kranendonk (1957, 1958) overlap and quadrupole interactions are together responsible for the intensity of the fundamental vibration band of hydrogen. Although both types of interaction should contribute to the intensity of the rotational band, it might be expected that in some cases the overlap effect would be relatively much smaller than the quadrupole effect; the overlap intensity is, moreover, difficult to calculate. On the other hand, the intensity of purely rotational transitions due to quadrupole interaction is easily deducible from the theory of Van Kranendonk; such calculations have been carried out by Colpa (1957) and Kiss (1957). In the following an outline of the theory will be presented.

The binary absorption coefficient of the pure gas, $\tilde{\alpha}_1$, defined by the equation,

$$\int (\alpha/\nu) d\nu = \int \tilde{\alpha} d\nu = \tilde{\alpha}_1 n_{H_2}^2,$$

is given by the general expression,

$$\tilde{\alpha}_1 = \frac{8\pi^3}{3hc} \sum_{i,f} P_i \frac{1}{2} \int |\mathfrak{y}_{if}(\mathbf{R})|^2 g(\mathbf{R})\, d\mathbf{R}. \tag{1}$$

Here P_i indicates the normalized Boltzmann factor for the initial state, i; $\mathfrak{y}_{if}(\mathbf{R})$ is the matrix element of the dipole transition, $i \rightarrow f$, and is dependent on the intermolecular distance, $\mathbf{R}$; $g(\mathbf{R})$ is the pair distribution function. For purely quadrupolar interaction $\mathfrak{y}_{if}$ has the form

$$\mathfrak{y}_{if}^{\text{quad}} = \sum_{m} \psi^*_{J_1'm_1'} \psi^*_{J_2'm_2'} \mathfrak{y}(\omega_1, \omega_2)\, \psi_{J_1m_1} \psi_{J_2m_2}\, d\omega_1\, d\omega_2 \tag{2}$$

where $\boldsymbol{\mu}(\boldsymbol{\omega}_1, \boldsymbol{\omega}_2)$, the resultant dipole moment induced by the molecular quadrupoles, is a function of the orientations, $\boldsymbol{\omega}_1$ and $\boldsymbol{\omega}_2$, of the molecules 1 and 2, and ψ_{Jm} is the eigenfunction of the rotational state J with the magnetic quantum number m. The dipole moment induced by the quadrupole moment, Q_1, of molecule 1 in molecule 2, which has the polarizability, α_2, is

$$\boldsymbol{\mu} = (\alpha_2 Q_1/R^4) f(\boldsymbol{\omega}_1),$$

where $f(\boldsymbol{\omega}_1)$ is a function of the orientation of molecule 1, and α_2 is assumed to be spherically symmetric. There is also a dipole moment of similar form induced in molecule 1 by the quadrupole moment of molecule 2.

When $\boldsymbol{\mu}$ is expanded in a series of spherical harmonics and a summation is carried out over the substates, m, equation (1) becomes

$$(3) \quad \tilde{\alpha}_1^{\text{quad}}(J) = \frac{48\pi^4}{h} \alpha_{H_2}{}^2 Q_{H_2}{}^2 \{P(J) - P(J+2)\} \times \frac{(J+1)(J+2)}{(2J+3)} \int_0^\infty e^{-V_0(R)/kT} \frac{dR}{R^6}$$

for the rotational transition, $J \rightarrow J+2$. $P(J)$ is the probability of finding the molecule in the state, J, so that $\sum_J (2J+1)P(J) = 1$; the term $P(J+2)$ takes account of induced emission. $V_0(R)$ is the Lennard–Jones intermolecular potential function,

$$V_0(R) = 4\epsilon \left\{ \left(\frac{\sigma}{R}\right)^{12} - \left(\frac{\sigma}{R}\right)^6 \right\}.$$

The total absorption coefficient is $\tilde{\alpha}_1(\text{quad.}) = \sum_J \tilde{\alpha}_1^{\text{quad}}(J)$, where $J = 0, 1, 2$, and 3.

For the enhancement of the rotational absorption by foreign gases equation (3) can be suitably modified. For the intermolecular potential function of two molecular types, a and b, the usual approximations were used:

$$\epsilon = \sqrt{(\epsilon_a \epsilon_b)}, \quad \sigma = (1/2)(\sigma_a + \sigma_b).$$

The values of the integral,

$$(4) \qquad I_T = \int_0^\infty e^{-V_0(R)/kT} \frac{dR}{R^6},$$

were obtained by numerical integration. For the quadrupole moment of hydrogen the theoretical value of James and Coolidge (1938), $Q_{H_2} = 0.45ea_0^2$, was used. Values of ϵ/k and σ of the pure gases were taken from the summary of Hirschfelder, Curtiss, and Bird (1954). Since polarizabilities are frequency-dependent, values for use in the calculations were determined from refractive indices at 400 cm^{-1}, obtained by extrapolation of refractive index data available in the literature for the visible and near infrared regions; for the gases used it was found that a plot of $1/(n-1)$ vs. ν^2 gave a well-defined straight line. The values of the constants used in the calculations, along with the values of I, are summarized in Table II; values of $\tilde{\alpha}_1(\text{quad.})$ are given in column 6 of Table III.

TABLE II

Data used in the calculations

Perturbing gas	α cm$^3\times10^{24}$	Potential constants for the H_2 – foreign gas pair ϵ/k, ° K	σ, Å	$I_{300°\,K}$ cm$^{-5}\times10^{-37}$	$I_{195°\,K}$ cm$^{-5}\times10^{-37}$	$I_{85°\,K}$ cm$^{-5}\times10^{-37}$
H_2	0.806	37.0	2.92	1.69	1.46	1.38
N_2	1.75	59.4	3.32	0.862	0.831	—
He	0.205	14.9	2.77	2.56	2.13	—
Ne	0.396	36.0	2.86	1.83	1.61	—
A	1.65	66.5	3.12	1.06	1.04	—
Kr	2.45	70.6	3.26	0.896	—	—
Xe	4.05	90.5	3.51	0.665	—	—

TABLE III

Experimental and calculated values of the absorption coefficients

T, ° K	Perturbing gas	Experimental α_1 cm^{-2} Amgt^{-2}	$\bar{\nu}$ cm^{-1}	$\tilde{\alpha}_1 = \alpha_1/\bar{\nu}$ sec^{-1} cm^6	Calculated $\tilde{\alpha}_1$ (quad.) sec^{-1} cm^6	$\tilde{\alpha}_1$(exptl.)/$\tilde{\alpha}_1$ (quad.)	$\tilde{\alpha}_1$ (exptl.) – $\tilde{\alpha}_1$ (quad.) sec^{-1} cm^6
300	n-H_2	2.06×10^{-2}	628	1.37×10^{-24}	1.09×10^{-24}	1.25	0.28×10^{-24}
	N_2	5.42 "	636	3.54 "	2.66 "	1.33	0.88 "
	He	0.63 "	602	0.44 "	0.108 "	4.0	0.33 "
	Ne	1.24 "	615	0.836 "	0.289 "	2.89	0.54 "
	A	6.71 "	635	4.41 "	2.94 "	1.50	1.5 "
	Kr	9.60 "	641	6.22 "	5.50 "	1.13	0.72 "
	Xe	17.4 "	642	11.1 "	11.0 "	1.01	0.1 "
195	n-H_2	1.76 "	613	1.19 "	1.06 "	1.12	0.13 "
	N_2	4.77 "	608	3.27 "	2.85 "	1.15	0.42 "
	He	0.27 "	590	0.187 "	0.0990 "	1.89	0.09 "
	Ne	0.82 "	592	0.475 "	0.284 "	1.67	0.19 "
	A	5.30 "	610	3.61 "	3.17 "	1.14	0.44 "
85	n-H_2	1.69 "	562	1.25 "	1.11 "	1.13	0.14 "
	e-H_2	1.63 "	521	1.30 "	1.23 "	1.06	0.07 "

DISCUSSION

The agreement of the experimental values of $\tilde{\alpha}_1$ and the calculated values of $\tilde{\alpha}_1$(quad.) is shown by the values of $\tilde{\alpha}_1$(exptl.)/$\tilde{\alpha}_1$(quad.) and $\tilde{\alpha}_1$(exptl.) $-\tilde{\alpha}_1$(quad.) given in the last two columns of Table III. In all cases $\tilde{\alpha}_1$(exptl.) is greater than $\tilde{\alpha}_1$(quad.). The results for 300° K form the most extensive series for the comparison of the data for the various perturbing molecules. Here the ratio, $\tilde{\alpha}_1$(exptl.)/$\tilde{\alpha}_1$(quad.) ranges from 4.0 to 1.02 and shows for the rare gases a monotonic dependence on the mass of the perturbing molecule. It was found that for all of the perturbing molecules $\tilde{\alpha}_1$(quad.) is a fairly smooth function of the molecular polarizability, α; thus, the graph of $\log_{10}\tilde{\alpha}_1$(quad.) vs. $\log_{10}(\alpha^2)$ in Fig. 9 is very nearly a straight line. In this type of graph the experimental values of $\tilde{\alpha}_1$ also form a smooth curve; however, this curve approaches that for $\tilde{\alpha}_1$(quad.) only for the heavier perturbing molecules. It will be noted that, although $\tilde{\alpha}_1$(exptl.) and $\tilde{\alpha}_1$(quad.) vary by factors of 25 and 100, respectively, $\tilde{\alpha}_1$(exptl.) $-\tilde{\alpha}_1$(quad.) changes by a factor of only 6 and, in the series of the rare gases, has its greatest value for argon.

There is thus a part of the experimental intensity which is not accounted for by the quadrupole induction of S transitions; this part of the intensity shows a smaller variation from molecule to molecule than the quadrupole intensity, and is important relative to the latter when the polarizability of the perturbing atom is small.

The theoretical and experimental investigations of the vibrational band show that in addition to the quadrupole induction of S transitions as calculated above, other factors might give rise to induced absorption in the low frequency spectrum of hydrogen. These are:

(1) The anisotropy of the polarizability of the hydrogen molecule. When the perturbing molecule has a quadrupole moment the anisotropy of the polarizability of the hydrogen molecule can give rise to some additional S-line intensity. This effect, which has been calculated by Colpa (1957), would be present among the gases used here only for pure hydrogen and hydrogen–nitrogen and would increase the intensity above $\tilde{\alpha}_1$(quad.) by only a few per cent.

(2) Transitions with $\Delta J = 0$ induced by quadrupole interaction. Transitions of this type have been predicted by the theory and are identified with the Q_Q component of the vibrational band. In the absence of any interaction with the relative translational motion of the colliding molecules the transition in the ground vibrational state of the hydrogen molecule would have zero frequency and would not give rise to absorption. The width of the Q_Q line shows that there is a relative kinetic energy effect, but it is probable that this would lead to absorption only at very low frequencies in the far infrared; the effect would therefore not be observed in the present experiments.

(3) Transitions with $\Delta J = 2$ due to interaction in the region of overlap forces. As pointed out above this contribution to the S-line intensity should be small and a more extended theoretical investigation would be necessary to assess its importance.

(4) Transitions with $\Delta J = 0$ induced by the overlap interaction. These transitions account for most of the intensity in the Q branch of the vibrational band. Experimentally it is found that the intensity is spread over two wide components, designated Q_P and Q_R, which have been interpreted as difference and summation tones of the vibrational frequency of the hydrogen molecule with a continuum of frequencies due to the relative motions of the molecules during the collision. If the same effect is present when no vibrational transition takes place, a continuum corresponding to the Q_R component should be superimposed on the rotational spectrum. Moreover, if the shape of this continuum resembles that of the Q_R component, the intensity should decrease rather rapidly towards higher frequencies in the region of observation, 300 to 1400 cm^{-1}. The physical process involved in the production of the continuum is the absorption of a photon by two molecules interacting in the region of overlap forces, the energy of the photon appearing after the collision as relative translational energy of the two molecules.

The shapes of the absorption profiles suggest that a "translational" absorption of this type actually exists as a continuum underlying the rotational

lines. In Fig. 10 the specific absorption profiles for pure hydrogen and for the enhancements by the rare gases have been plotted; for greater ease of comparison the ordinate scale has been adjusted so that the highest intensity, i.e. the maximum of the $S(1)$ line, has the same height in all cases. It is evident from these graphs that the shapes of the profiles are entirely consistent with the assumption of a continuum underlying the S lines and decreasing rapidly in intensity from 300 to 1400 cm^{-1}. The half width of the S lines and the intensity of the continuum vary greatly from molecule to molecule. The characteristic intensity distribution of the continuum affects

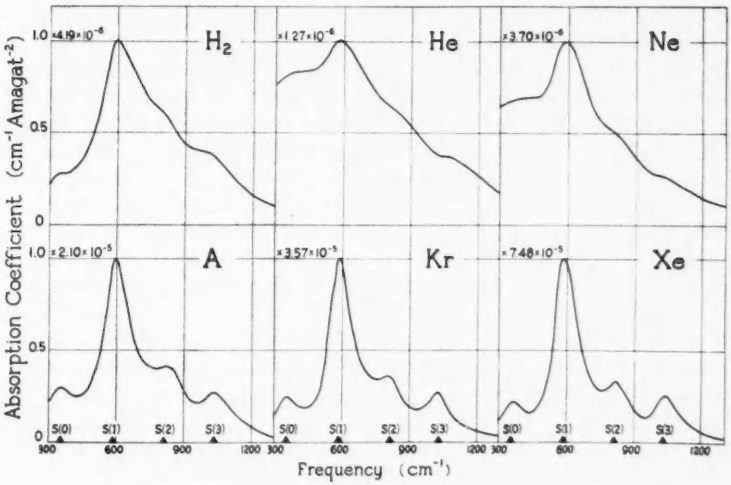


FIG. 10. Comparison of the specific absorption profiles of pure hydrogen and of the enhancement of the absorption in H_2 – rare gas mixtures.

particularly the intensity ratio of the maxima of the S lines, $I_{S(0)}{}^{max}/I_{S(1)}{}^{max}$. For xenon as the perturbing gas this ratio is 0.21 as compared with the value, 0.17, calculated on the basis of the quadrupole interaction; in this case the overlap continuum is very weak. For helium, on the other hand, the ratio is 0.82, indicating that the overlap interaction is responsible for a large fraction of the observed intensity. The other gas mixtures show translational absorption of intermediate intensity relative to the intensity of the S lines.

Observations on the vibrational band of hydrogen have shown that the intensity of the Q_R component, as well as its extent towards higher frequencies, decreases as the temperature is lowered. The intensity of the overlap continuum in the region of the rotational lines should therefore be less at lower temperatures and the observed absorption should be more nearly quadrupolar in origin. Table III shows that the ratio $\bar{\alpha}_1(\text{exptl.})/\bar{\alpha}_1(\text{quad.})$ is in fact more nearly equal to unity for a given perturbing molecule as the temperature is lowered.

The intensities of the rotational band for normal and equilibrium hydrogen at 85° K present some points of interest. The integrated absorption coefficient, α_1, is somewhat greater for normal hydrogen; however, both $\tilde{\alpha}_1$(exptl.) and $\tilde{\alpha}_1$(quad.) are greater for equilibrium hydrogen. The decrease in $\tilde{\alpha}_1$(exptl.) $-\tilde{\alpha}_1$(quad.) in going from normal to equilibrium hydrogen indicates a considerable decrease in the overlap intensity when the ortho–para ratio is decreased from 3.0 to 1.0. The reason appears to be the fact that paramolecules in the $J = 0$ state are spherically symmetric; a colliding pair of such molecules has therefore a center of symmetry and no induced dipole moment is produced.

In conclusion it can be stated with reasonable certainty that the pressure-induced absorption of hydrogen in the region 300–1400 cm^{-1} is due to both quadrupole and overlap interactions. When the perturbing molecule is heavy the absorption is produced predominantly by the polarization of the heavy molecule in the field of the quadrupole moment of the hydrogen molecule. For xenon as the perturbing gas the observed intensity is only 2% greater than that calculated for the quadrupole effect. This appears to be a striking confirmation of the accuracy of the value of the quadrupole moment of hydrogen used in the calculations, namely, the theoretical value of James and Coolidge. However, it must not be forgotten that the theory of the quadrupole effect as presented above uses approximations for the intermolecular potential of two molecular species; although these approximations are widely used, their validity is by no means established. Further elucidation of the observed absorptions requires the estimation of the overlap contribution to the S line intensity and a theory of the translational contribution. Some progress has already been made in analyzing empirically the observed asymmetric shape of the S lines with a view to separating the total intensity of the band into rotational and translational components; the details of this work will be reserved for a later communication.

ACKNOWLEDGMENT

The authors are greatly indebted to Professor J. Van Kranendonk for enlightening discussions on theoretical problems connected with pressure-induced absorption.

REFERENCES

BRITTON, F. R. and CRAWFORD, M. F. 1958. Can. J. Phys. **36**, 761.
CHISHOLM, D. A. and WELSH, H. L. 1954. Can. J. Phys. **32**, 291.
COLPA, J. P. 1957. Dissertation, Amsterdam.
COLPA, J. P. and KETELAAR, J. A. A. 1958. Mol. Phys. **1**, 14.
HARE, W. F. J. and WELSH, H. L. 1958. Can. J. Phys. **36**, 88.
HIRSCHFELDER, J. O., CURTISS, C. F., and BIRD, R. B. 1954. Molecular theory of gases and liquids (John Wiley & Sons, Inc., New York; Chapman and Hall, Ltd., London).
JAMES, H. M. and COOLIDGE, A. S. 1942. Astrophys. J. **87**, 438.
KETELAAR, J. A. A., COLPA, J. P., and HOOGE, F. N. 1955. J. Chem. Phys. **23**, 413.
KISS, Z. J. 1957. M.A. Thesis, University of Toronto, Toronto, Ontario.
STOICHEFF, B. P. 1957. Can. J. Phys. **35**, 730.
VAN KRANENDONK, J. 1952. Dissertation, Amsterdam.
——— 1957. Physica, **23**, 825.
——— 1958. Physica, **24**, 347.
VAN KRANENDONK, J. and BIRD, R. B. 1951. Physica, **17**, 953, 968.

NOTES

GAMMA RADIATION FROM THE $U^{235}(n, \gamma)U^{239}$ REACTION*

P. J. CAMPION, J. W. KNOWLES, G. MANNING,† AND G. A. BARTHOLOMEW

The γ-rays following thermal neutron bombardment of a target consisting of 200 g of uranium metal‡ with the U^{235} content depleted to approximately 0.034% were examined with a pair spectrometer (Kinsey and Bartholomew 1953) and a flat crystal diffraction spectrometer (Knowles 1959). With the former instrument, the region between 7.7 and 3.0 Mev was surveyed using a resolution of about 3%, and studied in more detail between 4.2 and 3.4 Mev with a resolution of 1%. The energy range from 3.4 to 0.14 Mev was examined with the crystal spectrometer, Laue diffracting from the (440) planes of a single germanium crystal, 25 mm thick. With this single crystal arrangement the resolution was 4% at 1 Mev. The region near 0.6 Mev, which contained a number of prominent γ-rays, was further studied with a resolution of about 0.5% by a double crystal arrangement, Laue diffracting from the (211) planes of calcite crystals, each 23 mm thick.

The spectra uncorrected for spectrometer efficiencies and γ-ray absorption are shown in Fig. 1 for the region above 3.0 Mev and in Fig. 2 for the region in the neighborhood of 0.6 Mev. The (smoothed) background spectrum, taken with an empty sample container, is shown as an emphasized broken line in Fig. 1. The spectrum in this figure was obtained with a constant line width of 150 kev, while that in the inset was taken with a line width of about 40 kev. At this resolution the γ-rays at 3.576 and 3.662 Mev are clearly separated. From the shape of the peak at 4.062 Mev under both resolutions it appears that there may be some significant unresolved radiation in the low energy tail of this peak. The energies and absolute intensities of the resolved γ-rays are given in Table I. The absolute intensities for γ-rays above 3.4 Mev were determined by the nickel comparison method (Kinsey and Bartholomew 1953). For radiations below 3.4 Mev the absolute intensities were found by direct comparison with the γ-rays of known intensity following the β-decay of Np^{239}, which, after several days irradiation, is in equilibrium with the (n, γ) reaction. The accuracy of the absolute intensities quoted in Table I is between 30 and 50%. The spectrum corrected for spectrometer efficiencies and γ-ray absorption is shown in Fig. 3. In this figure the ordinate is expressed in units of energy radiated per capture per Mev, and the cross-hatched area indicates the uncertainty in the absolute intensity measurements. The γ-rays following the β-decay of Np^{239} are well established (Hollander *et al.* 1956; Ewan *et al.*

*Issued as A.E.C.L. No. 753.

†Now at Norman Bridge Laboratory of Physics, California Institute of Technology, Pasadena, California.

‡We are indebted to the United States Atomic Energy Commission for providing this material.

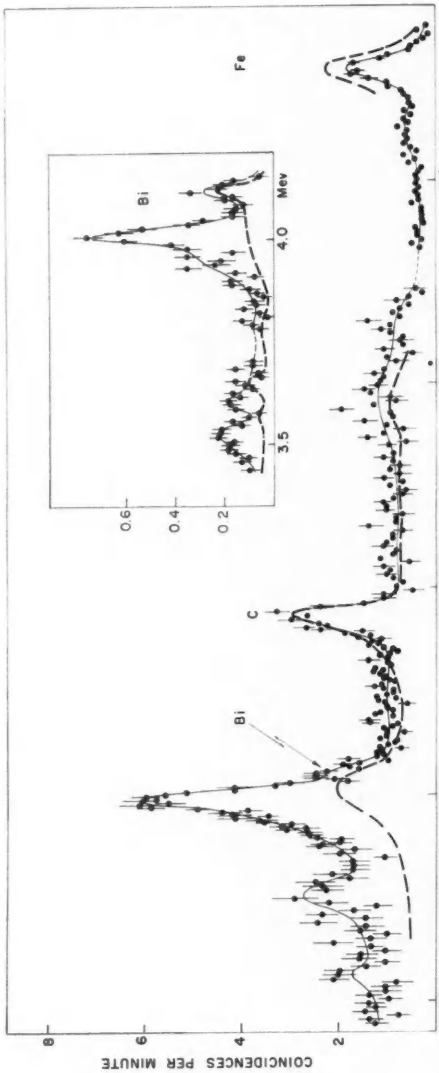


FIG. 1. The uncorrected γ-ray spectrum above 3.0 Mev following neutron bombardment of a target of separated U^{238}. The measurements were made with the pair spectrometer operating at a constant line width of 150 kev. The errors have been omitted from some points for clarity. The inset shows the data obtained in the region of 4 Mev using a line width of about 40 kev. In both cases the broken line indicates the background spectrum obtained with an empty sample container. The γ-rays at 4.2, 4.95, and 7.6 Mev are produced by neutron capture in a bismuth plug, in the graphite walls of the target container, and in the iron walls of the experimental hole respectively.

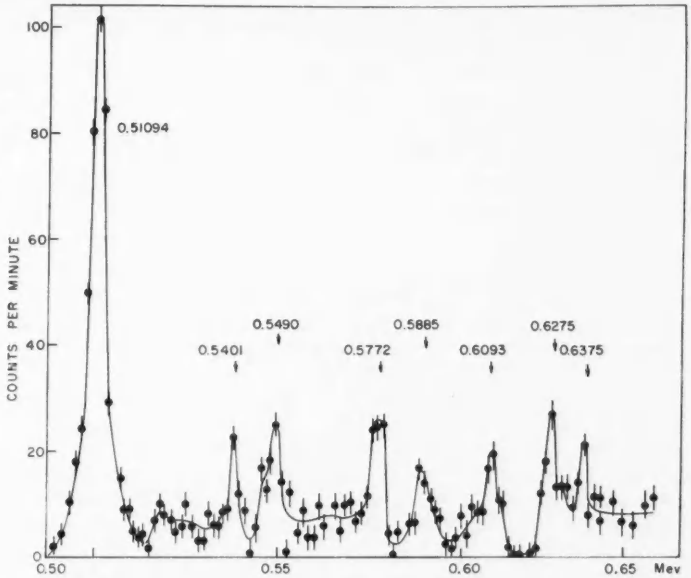


FIG. 2. The uncorrected γ-ray spectrum between 0.5 and 0.65 Mev following neutron bombardment of a target of separated U^{238}. The measurements were made with the double flat crystal diffraction spectrometer operating at a resolution of 0.5%. The γ-ray energies quoted are measured relative to the 0.51094 Mev annihilation radiation.

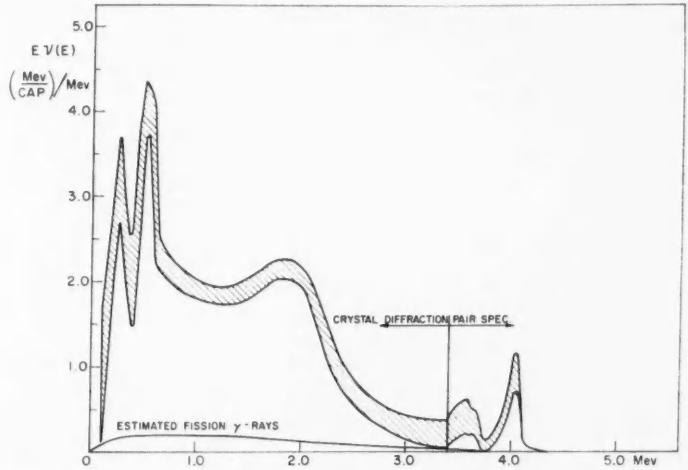


FIG. 3. The smoothed γ-ray spectrum following neutron bombardment of a target of separated U^{238}. The spectrum has been corrected for spectrometer efficiencies and γ-ray absorption. The γ-rays associated with the β-decay of Np^{239} have been subtracted. The estimated contribution due to prompt fission γ-rays is also shown.

1957) and these γ-rays have been subtracted from the curve of Fig. 3. The contribution to the remaining spectrum from prompt fission γ-rays from the residual U^{235} in the target was estimated from the data of Francis and Gamble (1955) and is indicated in Fig. 3. Owing to the difficulty in evaluating the contribution of fission product γ-rays no attempt has been made to portray the latter; however, current evidence seems to suggest that both the energy emitted per fission and the distribution in energy of these radiations are similar to those for the prompt fission γ-rays (Maienschein *et al.* 1958).

TABLE I

Resolved neutron capture gamma radiation from the $U^{238}(n,\gamma)U^{239}$ reaction

Energy, Mev	Intensity, γ-rays per capture
0.5401±0.0004	0.09
0.5490±0.0004	0.09
0.5772±0.0004	0.09
0.5885±0.0004	0.05
0.6093±0.0004	0.06
0.6275±0.0004	0.09
0.6375±0.0004	0.03
3.576±0.010	0.02
3.662±0.018	0.02
4.062±0.010	0.07

The energy radiated per capture, $\int_{0.14}^{\infty} E\nu(E)dE$, in the $U^{238}(n,\gamma)U^{239}$ reaction was found to be 4.3 Mev, which, in view of the probable errors in the intensity determinations, is in reasonable agreement with the neutron binding energy of U^{239}. For the latter quantity the value 4.63±0.15 Mev has been obtained from the $U^{238}(d,p)U^{239}$ reaction (Harvey 1951) and 4.78 Mev from closed decay energy cycles (Hyde and Seaborg 1957). No γ-rays which could be definitely ascribed to the $U^{238}(n,\gamma)U^{239}$ reaction were observed above 4.062 Mev, in contrast to the results of Kenney and Mattingly (1956). It is estimated that if uranium capture γ-rays are present in the energy region above 4.062 Mev their intensities must be less than 0.01 per capture.

An attempt was made to detect γ-rays in coincidence with the 4.062 Mev radiation. For this experiment a 23-g target of U^{238} metal was bombarded by a beam of thermal neutrons and the capture γ-rays were detected with two NaI(Tl) scintillation spectrometers (Manning and Bartholomew 1959). A differential discriminator associated with one of the detectors was set to accept pulses from the 4.062 Mev line and the spectrum in the other detector coincident with these pulses was observed with a 100-channel pulse height analyzer. The resulting spectrum, although not resolved into separate γ-rays, indicated that the 4.062 Mev radiation is not the ground state transition. The coincidence spectrum was consistent with a binding energy of about 4.7 Mev.

EWAN, G. T., KNOWLES, J. W., and MACKENZIE, D. R. 1957. Phys. Rev. **108**, 1308.
FRANCIS, J. E. and GAMBLE, R. L. 1955. Oak Ridge National Laboratory ORNL 1879, 20.
HARVEY, J. A. 1951. Phys. Rev. **81**, 353.
HOLLANDER, J. M., SMITH, W. G., and MIHELICH, J. W. 1956. Phys. Rev. **102**, 740.

HYDE, E. K. and SEABORG, G. T. 1957. Encyclopedia of Physics, Vol. XLII (Springer-Verlag, Berlin), p. 205.
KENNEY, R. W. and MATTINGLY, J. T. 1956. Bull. Am. Phys. Soc. **1** (8), 389.
KINSEY, B. B. and BARTHOLOMEW, G. A. 1953. Can. J. Phys. **31**, 537.
KNOWLES, J. W. To be published.
MAIENSCHEIN, F. C., PEELLE, R. W., ZOBEL, W., and LOVE, T. A. 1958. Second United Nations International Conference on the Peaceful Uses of Atomic Energy. Paper P/670.
MANNING, G. and BARTHOLOMEW, G. A. To be published.

RECEIVED NOVEMBER 26, 1958.
GENERAL PHYSICS BRANCH,
ATOMIC ENERGY OF CANADA LIMITED,
CHALK RIVER, ONTARIO.

THE EFFECT OF CONCENTRATION ON THE MAGNETIC SUSCEPTIBILITY OF TRAPPED ELECTRONS AND HOLES IN SEMICONDUCTORS

F. T. HEDGCOCK

In order to avoid the complications and difficulties inherent in the theory of impurity band formation in semiconductors, Hedgcock (1955) and Sonder and Stevens (1958) have tentatively proposed a simpler model to explain the anomalous magnetic susceptibility of certain impurity semiconductors at low temperatures. The idea underlying this treatment may perhaps be best explained by the following analogy proposed by Mott (1940). We may think of an electron bound in an impurity site as similar to an electron bound to a hydrogen nucleus but moving in a medium of dielectric constant equal to that of the material and with an appropriate effective mass. In practice this means that the radius of the smallest Bohr orbit is relatively large (in germanium about 48 Å). If now the concentration of occupied impurity centers becomes large enough, some of these orbits will overlap and it seems reasonable to assume the formation of impurity "molecules" in which two impurity ions share a pair of electrons in a manner analogous to the hydrogen molecule. The effect on the susceptibility would then be as follows: instead of the two electron spins contributing independently to a Langevin type of paramagnetism, they will now be aligned antiparallel and produce no resultant spin magnetic moment. This type of effect is what is needed to explain why at high bound carrier concentrations and at low temperatures the specimens show a greater diamagnetism than would be expected on the basis of an independent electron spin model. However, we wish now to show that although this concept is qualitatively attractive, it is quite inadequate quantitatively. For this purpose we require measurements of the magnetic susceptibility of semiconductors with suitable impurity concentrations. By "suitable" we mean (*a*) of high enough concentrations to have appreciable overlapping of bound electrons, (*b*) of concentrations low enough to retain a non-zero activation energy. Measurements of the low temperature magnetic properties of some doped germanium samples were made by the author (Hedgcock 1957) but it was found impossible to satisfy both conditions (*a*) and (*b*) in germanium.

In silicon, however, suitable concentrations can be obtained and measurements on such samples have been made by Sonder and Stevens (1958). From their measurements these authors deduce the number of electrons bound at low temperatures (from the measured Hall coefficients) and the number of magnetic centers at low temperatures from the magnetic susceptibility versus T^{-1} graphs. These deduced values are listed for two samples of high concentrations in Table I, columns (1) and (2).

TABLE I

(1) Charge carrier concentration deduced from Hall coefficient at high temp.	(2) No. of bound electrons derived from low temp. sus.	(3) No. of overlapping bound electrons (experimental)	(4) No. of overlapping bound electrons (calculated)
5×10^{18}	1.2×10^{18}	3.8×10^{18}	10^{18}
1.2×10^{18}	1.1×10^{18}	10^{17}	10^{16}

Column (3), which is the difference between (1) and (2), is interpreted as the number of overlapping bound electrons, i.e. the number of bound electrons which do not appear to contribute to the resultant magnetic moment.

We now wish to compare this number (column (3)) with the number of overlapping bound electrons which would be calculated on the basis of the model outlined above. We assume that when two impurity centers lie within a certain critical volume determined by the minimum potential energy of the system, a "molecular" impurity site is formed. The separation of the impurity centers for molecular behavior is found by modifying the energy expression given for molecular hydrogen (see Glasstone 1949) by including an effective mass and a dielectric constant appropriate to the material (here, silicon). For a perfectly random binary solid solution we can calculate the probability of finding two impurity centers close enough for overlapping to occur at any specified concentration. This can be done by following the simple statistical theory of solid solutions outlined by Smoluchowski (1951); column (4), Table I, shows the result of this calculation.*

It is seen that the agreement between columns (3) and (4), which compare the experimentally derived and theoretically estimated number of overlapping bound electrons, is only qualitatively correct. The simple model presented here is thus clearly inadequate. It seems therefore that to explain the large decrease in the paramagnetism of trapped charge carriers at low temperatures one must assume the existence of an impurity band. It should be noted, however, that attempts to interpret the results on germanium (Hedgcock 1957) on the basis of Mooser's (1955) theoretical treatment of the susceptibility due to trapped electrons in an impurity band were not conclusive. Moreover Sonder and Stevens (1958) were unable to explain their results in silicon on this basis although their argument is not altogether convincing. It seems

*It is perhaps of interest to note that the same calculation shows that complete molecular behavior of electronic bound states should be observed when the bound electron concentration reaches 7.5×10^{18}/cc in germanium and 4.3×10^{19}/cc in silicon for *n*-type materials.

therefore that Mooser's treatment should be extended to a more general model of an impurity band; however, Kohn (1957) has emphasized the difficulties inherent in such a calculation.

The author wishes to mention the generous support of this research given by the National Research Council of Canada and the Defence Research Board.

GLASSTONE, S. 1949. Theoretical chemistry (D. Van Nostrand Company, Inc.), p. 79.
HEDGCOCK, F. T. 1955. Susceptibility of Ge.Sb alloys. Ph.D. Thesis, University of Western Ontario, London, Ontario.
——— 1957. J. Electronics, **2**, 513.
KOHN, W. 1957. Solid state physics, Vol. 5 (Academic Press, Inc.).
MOOSER, E. 1955. Phys. Rev. **100**, 1589.
MOTT, N. F. and GURNEY, R. W. 1940. Electronic processes in ionic crystals (Oxford University Press), Chap. 3.
SMOLUCHOWSKI, R. 1951. Phys. Rev. **84**, 511.
SONDER, E. and STEVENS, D. K. 1958. Phys. Rev. To be published.

RECEIVED DECEMBER 5, 1958.
DEPARTMENT OF PHYSICS,
UNIVERSITY OF OTTAWA,
OTTAWA, ONTARIO.

LETTERS TO THE EDITOR

Under this heading brief reports of important discoveries in physics may be published. These reports should not exceed 600 words and, for any issue, should be submitted not later than six weeks previous to the first day of the month of issue. No proof will be sent to the authors.

Impurity molecule-ions and impurity molecules

The possibility of impurity molecule formation has been pointed out by Hedgcock (1959). It is conceivable that at intermediate impurity concentrations the bound impurity atoms do not overlap sufficiently to form bands, but that by pure chance, it is possible that a small number of impurity atoms combine to form impurity molecules. There is some experimental evidence for this point of view (Fletcher *et al.* 1954; Slichter 1955; Hedgcock 1959). It does not seem to have been pointed out that one might also expect the formation of impurity molecule-ions in heavily compensated samples, even when the *n*- or *p*-type excess is small.

We are investigating the formation of such impurity states theoretically using the effective mass approximation. Although atomic impurity states bear a close resemblance to the states of atomic hydrogen, this analogy does not hold for molecular impurity systems. The reason is that the "nuclei" of an impurity molecule are held rigidly in place by valence bonds with neighboring atoms. This fact has two interesting consequences.

(1) The rotational bands found in ordinary molecular spectra are absent in molecular impurity spectra.

(2) The internuclear distance is not determined by the minimum in the electronic energy curve, but is fixed by the impurity concentration.

Because of (2), the impurity molecules (or impurity molecule-ions) do not have a definite internuclear separation, and we expect the ionization energy to be dependent on concentration.

The analysis for molecule-ions is somewhat simpler theoretically than the corresponding calculation for impurity molecules. For a system of N acceptor impurities and $N+1$ donor impurities the matrix element of the perturbation in the Bloch representation is:

$$(n\mathbf{k}|U|n'\mathbf{k}') = \frac{-4\pi e^2}{\kappa V(\mathbf{k}-\mathbf{k}')^2}\left[1+\sum_{n=1}^{N} e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}_n^+} - \sum_{n=1}^{N} e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}_n^-}\right]$$

where κ is the dielectric constant and $\mathbf{r}_n^+$, $\mathbf{r}_n^-$ are the position vectors of the donor and acceptor impurity ions respectively. If we consider only the two positively charged ions (all other impurity ions being assumed far removed) this matrix element may be shown to become:

$$(n\mathbf{k}|U|n'\mathbf{k}') = \frac{-8\pi e^2}{\kappa V}\cdot\frac{\cos(\mathbf{k}-\mathbf{k}')\cdot a}{(\mathbf{k}-\mathbf{k}')^2}$$

where $2a$ is the vector from one impurity ion to the other. The above problem is now mathematically analogous to the double-well problem with a potential

$$V(r) = \frac{-e^2}{\kappa}\left[\frac{1}{|\mathbf{r}+\mathbf{a}|}+\frac{1}{|\mathbf{r}-\mathbf{a}|}\right].$$

If the effective mass is isotropic, an exact solution is possible, otherwise a variational approach must be adopted.

Detailed results will be published at a later date.

FLETCHER, R. C., YAGER, W. A., PEARSON, G. L., and MERRIT, F. R. 1954. Phys. Rev. **95**, 844.
HEDGCOCK, F. T. 1959. Can. J. Phys. **37**, 381.
SLICHTER, C. 1955. Phys. Rev. **99**, 479.

RECEIVED DECEMBER 15, 1958.

DEPARTMENT OF PHYSICS,
MCMASTER UNIVERSITY,
HAMILTON, ONTARIO.

R. R. HAERING

NOTES TO CONTRIBUTORS

Canadian Journal of Physics

MANUSCRIPTS

General.—Manuscripts, in English or French, should be typewritten, double spaced, on paper 8½×11 in. **The original and one copy are to be submitted.** Tables and captions for the figures should be placed at the end of the manuscript. Every sheet of the manuscript should be numbered. Style, arrangement, spelling, and abbreviations should conform to the usage of recent numbers of this journal. Greek letters or unusual signs should be written plainly or explained by marginal notes. Characters to be set in boldface type should be indicated by a wavy line below each character. Superscripts and subscripts must be legible and carefully placed. Manuscripts and illustrations should be carefully checked before they are submitted. Authors will be charged for unnecessary deviations from the usual format and for changes made in the proof that are considered excessive or unnecessary.

Abstract.—An abstract of not more than about 200 words, indicating the scope of the work and the principal findings, is required, except in Notes.

References.—References should be listed **alphabetically by authors' names,** unnumbered, and typed after the text. The form of the citations should be that used in current issues of this journal; in references to papers in periodicals, titles should not be given and only initial page numbers are required. The names of periodicals should be abbreviated in the form given in the most recent *List of Periodicals Abstracted by Chemical Abstracts.* All citations should be checked with the original articles and each one referred to in the text by the authors' names and the year.

Tables.—Tables should be numbered in roman numerals and each table referred to in the text. Titles should always be given but should be brief; column headings should be brief and descriptive matter in the tables confined to a minimum. Vertical rules should not be used. Numerous small tables should be avoided.

ILLUSTRATIONS

General.—All figures (including each figure of the plates) should be numbered consecutively from 1 up, in arabic numerals, and each figure referred to in the text. The author's name, title of the paper, and figure number should be written in the lower left corner of the sheets on which the illustrations appear. Captions should not be written on the illustrations.

Line drawings.—Drawings should be carefully made with India ink on white drawing paper, blue tracing linen, or co-ordinate paper ruled in blue only; any co-ordinate lines that are to appear in the reproduction should be ruled in black ink. Paper ruled in green, yellow, or red should not be used. All lines must be of sufficient thickness to reproduce well. Decimal points, periods, and stippled dots must be solid black circles large enough to be reduced if necessary. Letters and numerals should be neatly made, preferably with a stencil (**do NOT use typewriting**) and be of such size that the smallest lettering will be not less than 1 mm high when the figure is reduced to a suitable size. Many drawings are made too large; originals should not be more than 2 or 3 times the size of the desired reproduction. Whenever possible two or more drawings should be grouped to reduce the number of cuts required. In such groups of drawings, or in large drawings, full use of the space available should be made; the ratio of height to width should conform to that of a journal page (4¾×7½ in.), but allowance must be made for the captions. **The original drawings and one set of clear copies (e.g. small photographs) are to be submitted.**

Photographs.—Prints should be made on glossy paper, with strong contrasts. They should be trimmed so that essential features only are shown and mounted carefully, with rubber cement, on white cardboard, with no space between those arranged in groups. In mounting, full use of the space available should be made. **Photographs are to be submitted in duplicate;** if they are to be reproduced in groups one set should be mounted, the duplicate set unmounted.

REPRINTS

A total of 50 reprints of each paper, without covers, are supplied free. Additional reprints, with or without covers, may be purchased at the time of publication.

Charges for reprints are based on the number of printed pages, which may be calculated approximately by multiplying by 0.6 the number of manuscript pages (double-spaced typewritten sheets, 8½×11 in.) and including the space occupied by illustrations. Prices and instructions for ordering reprints are sent out with the galley proof.

Contents

C. D. Niven—A proposed mechanism for ice fraction - - - - - - - 247

K. W. Geiger and G. N. Whyte—Absolute standardization of radioactive neutron sources. I. Activation of manganese bath - - - - - - - - 256

H. E. Shanks and R. W. Bickmore—Four-dimensional electromagnetic radiators - 263

A. E. Scheidegger—On the theory of flow of underground fluids in compressible strata - 276

B. A. McIntosh—An experimental study of interception noise in electron streams at microwave frequencies - - - - - - - - - - - - 285

J. P. Hobson—First adsorbed layer of helium at 4.2° K - - - - - - - 300

D. R. Hay and G. E. Poaps—Prolonged signal fade-out on a short microwave path - 313

A. G. W. Cameron—Multiple neutron capture in the Mike fusion explosion 322

W. A. Little—The transport of heat between dissimilar solids at low temperatures - 334

D. D. Betts—Solid harmonics as basis functions for cubic crystals - - - 350

Z. J. Kiss, H. P. Gush, and H. L. Welsh—The pressure-induced rotational absorption spectrum of hydrogen. I. A study of the absorption intensities 362

Notes:

P. J. Campion, J. W. Knowles, G. Manning, and G. A. Bartholomew—Gamma radiation from the $U^{238}(n, \gamma)U^{239}$ reaction - - - - - - - - - - 377

F. T. Hedgcock—The effect of concentration on the magnetic susceptibility of trapped electrons and holes in semiconductors - - - - - - 381

Letters to the Editor:

R. R. Haering—Impurity molecule-ions and impurity molecules - - - 384

Printed in Canada *by* University of Toronto Press